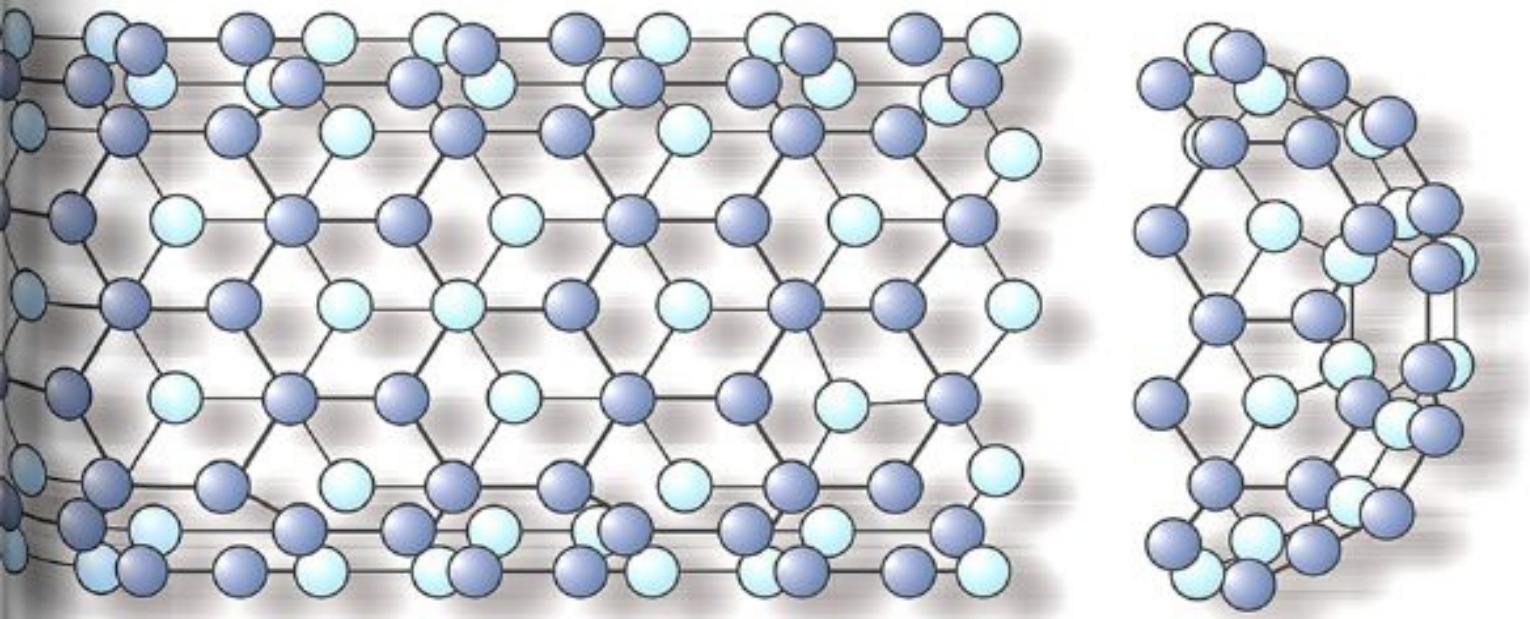


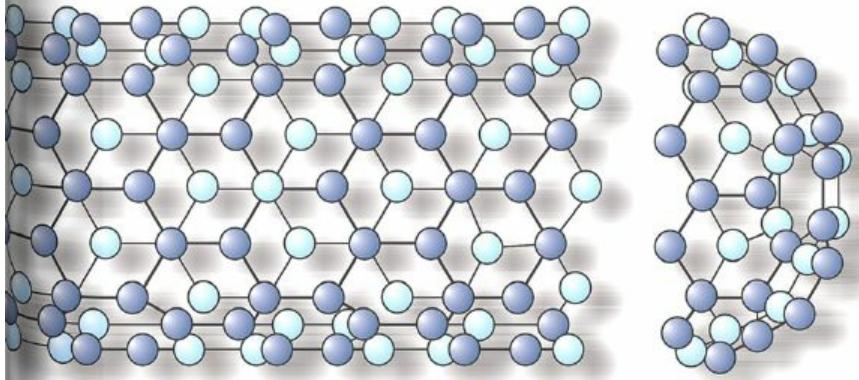
A TEXTBOOK OF
**APPLIED
PHYSICS**
SECOND EDITION
VOLUME II



A.K. JHA



A TEXTBOOK OF
**APPLIED
PHYSICS**
SECOND EDITION
VOLUME II



A.K. JHA



A Textbook of Applied Physics

Volume II
Second Edition

Dr. A.K. Jha

M.Sc. (DU), Ph.D. (IITD)

Department of Applied Physics

Delhi Technological University

Delhi-110042

Presently

Professor & Head

Applied Sciences Department

Ambedkar Institute of Advanced Communication

Technologies & Research, (GGSIPU)

Geeta Colony, Delhi-110031



I.K. International Publishing House Pvt. Ltd.

NEW DELHI

Published by

I.K. International Publishing House Pvt. Ltd.
S-25, Green Park Extension
Uphaar Cinema Market
New Delhi-110 016 (India)
E-mail: info@ikinternational.com
Website: www.ikbooks.com

ISBN 978-93-81141-76-2

10 9 8 7 6 5 4 3 2 1

© 2012 I.K. International Publishing House Pvt. Ltd.

Reprint 2015

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or any means: electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission from the publisher.

Published by Krishan Makhijani for I.K. International Publishing House Pvt. Ltd., S-25, Green Park Extension, Uphaar Cinema Market, New Delhi-110 016 and Printed by Rekha Printers Pvt. Ltd., Okhla Industrial Area, Phase II, New Delhi-110 020.

Preface

This gives me immense pleasure in bringing out the second edition of volume II of *Textbook of Applied Physics*. I hope that the book will be helpful to the students of B.E./B.Tech/B.Sc. and Diploma in Engineering. Due to the short duration of semester, the students need to have subject materials in a simple and easy to understand way. This is where the book is intended to serve the students. To further cover the syllabi of various other universities four new chapters, namely, *Physics of Semiconductors; Dielectric, Ferroelectric and Piezoelectric Properties of Materials; Superconductivity; and Nanomaterials* have been added in the present edition. This addition will make this book useful for Materials Science paper also. Despite our best efforts, some mistakes might have crept in. We shall be thankful to the students and teacher colleagues who kindly point out such mistakes to us. We wish reader students grand success and a bright future.

Dr. A. K. Jha
E-mail: prof.akjha@gmail.com

Contents

<i>Preface</i>	v
1. Quantum Physics	3
1.1 Introduction—Limitations of Classical Mechanics	3
1.2 Planck's Radiation Law—Quantum Theory	3
1.3 Photons	7
1.4 Interaction of Photons with Atoms	8
1.5 Photoelectric Effect	9
1.6 Einstein's Photoelectric Equation	12
1.7 Photoelectric Cells	14
1.8 Compton Effect	21
1.9 Pair Production and Annihilation	27
1.10 Dual Nature of Radiation: De-Broglie Waves-Matter Waves	28
1.11 De-Broglie Wavelength of Electron	29
1.12 Davisson and Germer Experiment—Verification of Matter Waves	30
1.13 Wave Packet: Group Velocity and Phase (or Wave) Velocity	32
1.14 Uncertainty Principle	35
1.15 Wave Function	41
1.16 Bra-Ket Notation in Quantum Mechanics	43
1.17 Expectation Values	44
1.18 Operators in Quantum Mechanics	45
1.19 Schrödinger Equation: Time Dependent Form	47
1.20 Schrödinger's Equation: Steady State Form (Time Independent)	48
1.21 Eigenvalues and Eigenfunctions	49
1.22 Degenerate and Non-degenerate Eigenfunctions	49
1.23 Particle in a Box	50
1.24 Simple Harmonic Oscillator	52
1.25 Rectangular Potential Barrier: Tunnel Effect	55
2. Statistical Mechanics	67
2.1 Introduction	67
2.2 Microscopic and Macroscopic Systems	67

2.3 Phase Space	68
2.4 Equal Probability Hypothesis	69
2.5 Statistical Equilibrium	69
2.6 Maxwell-Boltzmann Distribution Law	70
2.7 Maxwell's Speed Distribution Law	71
2.8 Law of Equipartition of Energy	74
2.9 Quantum Statistics	76
2.10 Bose-Einstein Statistics	76
2.11 Fermi-Dirac Statistics	79
2.12 Fermi Distribution Function-Fermi Energy	81
2.13 Energy of a Fermi-Dirac System	82
3. Nuclear Physics	89
3.1 Introduction	89
3.2 Structure of Nucleus	89
3.3 General Properties of Nucleus	90
3.4 Nuclear Force	93
3.5 Nuclear Binding Energy-Stability of Nuclei	96
3.6 Decay of Unstable Nuclei	100
3.7 The Radioactive Decay Law	101
3.8 Artificial (or Induced) Radioactivity	108
3.9 Applications of Radioactivity	119
3.10 Alpha Decay	110
3.11 Beta Decay	115
3.12 Gamma Decay	119
3.13 Nuclear Reactions	120
3.14 Nuclear Models	123
3.15 Nuclear Fission	128
3.16 Nuclear Reactor	131
3.17 Nuclear Fusion	133
3.18 Nuclear Holocaust	135
3.19 Accelerators	138
3.20 Linear Accelerator (LINAC)	140
3.21 Cyclotron	141
3.22 Betatron	144
3.23 Nuclear Radiation Detectors	147
3.24 Geiger-Muller (GM) Counter	149
3.25 Solid State Detectors	150
3.26 Scintillation Detectors	150
3.27 Cloud Chamber	151
3.28 Bubble Chamber	151
4. Electromagnetic Theory	157
4.1 Introduction	157
4.2 Motion of Charged Particles in Electric and Magnetic Force	157
4.3 Magnetic Focussing	160
4.4 Gauss's Law	161

4.5 Applications of Gauss's Law	163
4.6 Maxwell's Displacement Current	176
4.7 Continuity of Current	178
4.8 Maxwell's Equations	182
4.9 Wave Equation for Plane Electromagnetic Wave in Vacuum	183
4.10 Sinusoidal Electromagnetic Waves	186
4.11 Energy and Momentum in Electromagnetic Waves	188
4.12 Poynting's Theorem: Energy in Electromagnetic Waves	192
4.13 Electromagnetic Waves in a Dielectric Medium	196
4.14 Electromagnetic Waves in a Conducting Medium	198
4.15 The Electromagnetic Spectrum	201
5. X-rays Production and Properties & Ultrasonics	207
5.1 X-rays: Production and Properties	207
5.2 Ultrasonics and Acoustics of Buildings	217
6. Physics of Semiconductors	239
6.1 Formation of Energy Bands in Solids	239
6.2 Distinction between Metals, Insulators and Semiconductors	240
6.3 Current Carriers (Electrons and Holes) in Semiconductors	243
6.4 Intrinsic Semiconductors	245
6.5 Doping of Semiconductors: Extrinsic Semiconductors	245
6.6 Types of Extrinsic Semiconductors: N- and P-type Semiconductors	245
6.7 Comparison between Intrinsic and Extrinsic Semiconductor	248
6.8 Comparison between N- and P-type Semiconductors	248
6.9 Characteristics of a Semiconductor	249
6.10 Electrical Conductivity of Extrinsic Semiconductors	249
6.11 Carrier Concentration in Intrinsic Semiconductors	254
6.12 Electrical Conductivity: Variation with Temperature	258
6.13 Drift and Diffusion Currents: Einstein's Equation	260
6.14 Generation and Recombination of Minority Carriers	263
6.15 Hall Effect	264
6.16 Formation of p-n Junction Diode	270
6.17 Depletion Layer and Barrier Potential in a p-n Junction Diode	271
6.18 Biasing of p-n Junction Diode	272
6.19 Characteristics of a p-n Junction Diode	273
6.20 Zener Diode	277
6.21 Tunnel Diode	280
7. Dielectric, Ferroelectric and Piezoelectric Properties of Materials	286
7.1 Dielectric Materials: Introduction	286
7.2 Review of Basics	287
7.3 Dipole Moment and Polarization	288
7.4 Dielectric Constant	288
7.5 Dielectrics and Coulomb's law	291
7.6 Gauss's Law in Dielectrics	291
7.7 Electrical Susceptibility and Polarizability	293

7.8 Local Electric Field: Claussius-Mossotti Equation	297
7.9 Types of Polarization	302
7.10 Electronic Polarization	304
7.11 Dielectric Dispersion: Dielectric Loss	307
7.12 Dielectric Smart Materials	308
7.13 Piezoelectricity	308
7.14 Pyroelectricity	311
7.15 Ferroelectricity	311
7.16 Electrets	318
7.17 Ceramics	319
7.18 Electrostriction	319
7.19 Dielectric Strength/Breakdown	320
8. Superconductivity	325
8.1 Discovery of Superconductivity	325
8.2 Zero Resistance - Persistent Current	328
8.3 A.C. Resistivity in Superconductors	328
8.4 Effect of Magnetic Field: Critical Magnetic Field	329
8.5 Meissner Effect – Flux Exclusion	333
8.6 Type I and II Superconductors	335
8.7 Penetration Depth	337
8.8 Thermal Properties	338
8.9 Absorption of Electromagnetic Radiation	342
8.10 Isotope Effect	343
8.11 London Equation	344
8.12 BCS Theory	348
8.13 Magnetic Flux Quantization	353
8.14 Tunnelling of Single Particle	354
8.15 Josephson Effect-Cooper-Pair Tunnelling	355
8.16 Macroscopic Quantum Interference - SQUID	359
8.17 High-Temperature Superconductors	361
8.18 Superconducting Fullerenes	364
8.19 Applications of Superconductors	365
9. Nanomaterials	370
9.1 Introduction	370
9.2 Nano Science & Technology: Past and Future	371
9.3 Nanoparticles: Deviation from Bulk Behaviour	372
9.4 Metal Nanoclusters	373
9.5 Quantum Structures: Dimension Effects	375
9.6 Carbon Nanostructures	379
9.7 Other Applications of Nanomaterials	387
<i>Index</i>	395

Unit I

1

Quantum Physics ---

1.1 INTRODUCTION—LIMITATIONS OF CLASSICAL MECHANICS

We start studying mechanics based on Newton's laws of motion from early classes. The mechanics based on the Newton's laws, particularly second law of motion, is often referred to as classical mechanics or Newtonian mechanics. The classical mechanics successfully explains the motion of macroscopic bodies including terrestrial bodies. However, it fails to explain the motion of atomic particles. For example, if classical mechanics is applied to an atom, the electrons moving around the nucleus experience centripetal acceleration and hence must radiate energy in the form of electromagnetic waves. Therefore, the energy of the electrons should decrease continuously and ultimately they should collapse in the nucleus, contrary to the reality. The classical mechanics also failed to explain photoelectric effect, Compton effect, Raman effect, black body radiations, discrete nature of atomic spectra, etc. These problems could be overcome with the development of quantum mechanics or wave mechanics. The development of quantum mechanics revolutionized the different branches of physical sciences. Planck played an important role in establishing the quantum nature of radiation. Later Schrodinger, Heisenberg and other developed what we call quantum or wave mechanics. Not only physics but entire physical sciences such as chemistry, materials science, etc. are interpreted using quantum mechanics.

1.2 PLANCK'S RADIATION LAW—QUANTUM THEORY

It was observed from the black body radiation curves (energy E vs wavelength λ) that the energy is not uniformly distributed in radiation spectrum. At a given temperature, the intensity of radiation increases with increase of wavelength and becomes maximum at a particular wavelength (Fig. 1.1). Beyond this wavelength intensity of heat radiation decreases. An increase in temperature causes a decrease in λ_m such that $\lambda_m T = \text{const.} = 0.2896 \text{ cm K}$ (Wien's displacement law). Also, an increase in temperature causes an increase in energy emission for all wavelengths.

It was difficult to understand the black body radiations using the existing classical theory. In order to explain the experimentally observed distribution of energy in the spectrum of black body,

Planck in 1901 introduced the extremely important idea of *quantum theory of heat radiation*. According to Planck, energy is emitted in the form of packets or quanta called photons. Each photon has an energy $h\nu$, where h is the Planck's constant and ν is the frequency of radiation.

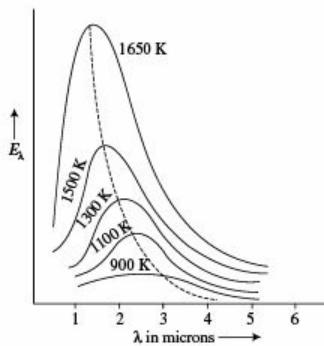


Fig. 1.1 Energy spectrum of black body radiations.

Plank derived this law by making the following assumptions.

- A chamber containing black body radiations contains simple harmonic oscillators or resonators of molecular dimensions (known as *Planck's oscillators or resonators*) which can vibrate with all possible frequencies.
- The frequency of radiation emitted by an oscillator is the same as the frequency of its vibration.
- An oscillator can emit (or absorb) energy in the multiples of a small unit called quantum (photon) and cannot emit (or absorb) energy in a continuous manner. An oscillator vibrating with frequency ν can emit (or absorb) energy in units or quanta of magnitude $h\nu$, i.e. an oscillator can emit (or absorb) only discrete energy values of $h\nu$, $2h\nu$, $3h\nu$,.... or $nh\nu$ where $n = 1, 2, 3, \dots$

In other words, exchange of energy between radiation and matter cannot take place continuously but only in discrete values, integral multiple of $h\nu$.

If N is the total number of Planck's oscillators and E their total energy, then average energy per Planck's oscillator is given by

$$\bar{\epsilon} = \frac{E}{N} \quad \dots(1.1)$$

According to Maxwell distribution law, if ϵ is a certain amount of energy, the probabilities that a system will have oscillator with energies $0, \epsilon, 2\epsilon, 3\epsilon, \dots$ are in the ratio

$$1 : \exp\left(\frac{-\epsilon}{kT}\right) : \exp\left(-\frac{2\epsilon}{kT}\right) : \exp\left(-\frac{3\epsilon}{kT}\right) : \dots$$

where k is Boltzmann constant and T the absolute temperature. If N_0 is the member of oscillators having zero energy, then the number of oscillators N_1 having ϵ energy will be $N_0 e^{-\epsilon/kT}$. Similarly, the number of oscillators N_2 having energy 2ϵ will be $N_0 e^{-2\epsilon/kT}$. In general, the number of oscillators N_r having $r\epsilon$ energy will be $N_0 e^{-r\epsilon/kT}$. Therefore, the total number of oscillators is given by

$$\begin{aligned} N &= N_0 + N_1 + N_2 + \dots + N_r + \dots \\ &= N_0 + N_0 e^{-\epsilon/kT} + N_0 e^{-2\epsilon/kT} + \dots + N_0 e^{-r\epsilon/kT} + \dots \\ &= N_0 [1 + e^{-\epsilon/kT} + e^{-2\epsilon/kT} + \dots + e^{-r\epsilon/kT} + \dots] \end{aligned}$$

Putting $e^{-\epsilon/kT} = y$, we get

$$N = N_0 [1 + y + y^2 + \dots + y^r + \dots] = \frac{N_0}{1-y} \quad \dots(1.2)$$

The total energy of Planck's oscillators is given by

$$\begin{aligned} E &= 0 \times N_0 + \epsilon \times N_1 + 2\epsilon \times N_2 + \dots + r\epsilon \times N_r + \dots \\ &= 0 + \epsilon N_0 e^{-\epsilon/kT} + 2\epsilon N_0 e^{-2\epsilon/kT} + \dots + r\epsilon N_0 e^{-r\epsilon/kT} + \dots \\ &= N_0 \epsilon [e^{-\epsilon/kT} + 2e^{-2\epsilon/kT} + \dots + re^{-r\epsilon/kT} + \dots] \\ &= N_0 \epsilon [y + 2y^2 + \dots + ry^r + \dots] \end{aligned}$$

Let

$$S = y + 2y^2 + \dots + ry^r + \dots$$

\therefore

$$Sy = y^2 + 2y^3 + \dots + (r-1)y^r + \dots$$

Subtracting, we get

$$S - Sy = y + y^2 + \dots + y^r + \dots$$

$$S(1-y) = \frac{y}{1-y}$$

or

$$S = \frac{y}{(1-y)^2}$$

$$E = N_0 \epsilon S = N_0 \epsilon \frac{y}{(1-y)^2}$$

Therefore, the average energy is given by

$$\begin{aligned} \bar{\epsilon} &= \frac{E}{N} = \frac{N_0 \epsilon \frac{y}{(1-y)^2}}{\frac{N_0}{1-y}} \\ &= \frac{\epsilon y}{(1-y)} = \frac{\epsilon e^{-\epsilon/kT}}{1 - e^{-\epsilon/kT}} \end{aligned}$$

or,

$$\bar{\epsilon} = \frac{\epsilon}{e^{\epsilon/kT} - 1}$$

According to Planck's assumption, $\epsilon = h\nu$, therefore

$$\bar{\epsilon} = \frac{h\nu}{e^{h\nu/kT} - 1} \quad \dots(1.3)$$

According to the classical theory average energy of a oscillator is kT . Thus, the average energy obtained using Planck's quantum theory is different from the classical theory.

It can be shown that the number of oscillators per unit volume (N) in frequency range ν and $\nu + d\nu$ is given by

$$N = \frac{8\pi\nu^2}{c^3} d\nu \quad \dots (1.4)$$

Thus, the total energy density per unit volume in the frequency range is given by the product of Eqs. (1.3) and (1.4)

$$E_\nu d\nu = \frac{8\pi h\nu^2}{c^3} d\nu \times \frac{h\nu}{e^{h\nu/kT} - 1}$$

or $E_\nu d\nu = \frac{8\pi h\nu^3}{c^3} \cdot \frac{1}{e^{h\nu/kT} - 1} d\nu \quad \dots(1.5)$

This is known as *Planck's radiation law*, in terms of frequency. This can also be expressed in terms of wavelength λ , as follows.

$$\begin{aligned} \nu &= \frac{c}{\lambda} \text{ and } d\nu = \left| -\frac{c}{\lambda^2} d\lambda \right| \\ E_\lambda d\lambda &= \frac{8\pi h}{c^3} \left(\frac{c^3}{\lambda^3} \right) \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \left| -\frac{c}{\lambda^2} d\lambda \right| \\ &= \frac{8\pi hc}{\lambda^5} \cdot \frac{d\lambda}{e^{\frac{hc}{\lambda kT}} - 1} \end{aligned} \quad \dots(1.6)$$

This expression agrees well with the experimental curve throughout the whole range of wavelengths. Wien's displacement and Rayleigh-Jeans law can be obtained from the Planck's radiation law.

Wien's displacement law: For shorter wavelengths, $e^{hc/\lambda kT}$ becomes large compared to unity ($e^{hc/\lambda kT} \gg 1$), hence from Eq. (1.6), we have

$$\begin{aligned} E_\lambda d\lambda &= \frac{8\pi hc}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} d\lambda \\ &= \frac{8\pi hc}{\lambda^5} e^{-\frac{hc}{\lambda kT}} d\lambda \end{aligned} \quad \dots(1.7)$$

which is Wien's displacement law.

Rayleigh-Jean's law: For longer wavelengths, $e^{hv/\lambda kT}$ can be approximated to $\left(1 + \frac{hc}{\lambda kT}\right)$ and hence Planck's law reduces to

$$\begin{aligned} E_\lambda d\lambda &= \frac{8\pi hc}{\lambda^5} \left(\frac{1}{1 + \frac{hc}{\lambda kT}} - 1 \right) d\lambda \\ &= \frac{8\pi hc}{\lambda^5} \cdot \frac{\lambda kT}{hc} d\lambda \\ &= \frac{8\pi kT}{\lambda^5} \cdot d\lambda \end{aligned} \quad \dots(1.8)$$

This is Rayleigh-Jean's law.

1.3 PHOTONS

A photon is a quantum (or a packet or a unit) of energy given out by a source of light. In other words, light consists of photons and its total energy is equal to the sum of the energies of individual photons.

The important characteristics of photons are given below.

- (i) Like light radiation, photons travel in a straight line.
- (ii) All photons emitted by a light source travel through space with speed of light and have frequency equal to that of light.
- (iii) The energy of a photon of frequency v is hv . That is,

$$\text{energy, } E = hv$$

where, h is Planck's constant.

Since, speed of light,

$$c = v\lambda \quad \text{or} \quad v = \frac{c}{\lambda}$$

$$E = \frac{hc}{\lambda} \quad \dots(1.9)$$

(iv) If m_0 is the rest mass of a particle, then its mass when moving with speed v is given by

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

For photon, $v = c$

$$\therefore m_0 = 0$$

That is, *rest mass of a photon is zero*. In other words, a photon cannot exist at rest.

- (v) The energy of a photon is usually expressed in electron volt (eV) ($1 \text{ eV} = 1.6 \times 10^{-19} \text{ J}$).
 (vi) A photon does not carry any charge, i.e., it is neutral.

Example 1.1

What is the frequency of a photon having energy $3.3 \times 10^{-20} \text{ J}$? ($\hbar = 6.6 \times 10^{-34} \text{ Js}$)

Solution.

$$\begin{aligned}\text{Here, } E &= 3.3 \times 10^{-20} \text{ J} \\ \text{Frequency, } v &=? \\ \text{Energy, } E &= \hbar v \\ \text{or, } v &= \frac{E}{\hbar} \\ &= \frac{3.3 \times 10^{-20} \text{ J}}{6.6 \times 10^{-34} \text{ Js}} \\ &= 0.5 \times 10^{14} \text{ Hz} \\ &= 5.0 \times 10^{13} \text{ Hz}\end{aligned}$$

Example 1.2

The energy of a photon is 3.45 eV. What is the wavelength associated with it?

Solution

$$\begin{aligned}\text{Here, } E &= 3.45 \text{ eV} \\ &= 3.45 \times 1.6 \times 10^{-19} \text{ J} \\ \text{As } E &= \hbar v = \frac{hc}{\lambda} \\ \lambda &= \frac{hc}{E} \\ &= \frac{6.6 \times 10^{-34} \times 3 \times 10^8}{3.45 \times 1.6 \times 10^{-19}} \\ &= \frac{19.8 \times 10^{-26}}{5.52 \times 10^{-19}} \\ &= 3.58 \times 10^{-7} \text{ m} \\ &= 3580 \text{ Å}\end{aligned}$$

1.4 INTERACTION OF PHOTONS WITH ATOMS

The nature of interaction of photons with the orbital electrons in the atom of the matter depends on the perpendicular distance between the direction of incident photon and the centre of the nucleus of

the atom. This distance is known as the interaction parameter, ‘ b ’ and is shown in Fig. 1.2. Depending on the relationship between b and the atomic radius (R), any of the following three physical phenomena can take place.

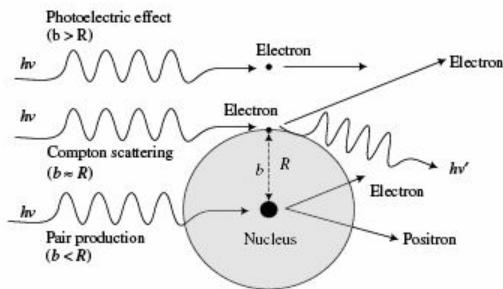


Fig. 1.2 Photoelectric effect, compton effect and pair production in interactions of photons with matter.

If $b > R$, the electron is ejected by the photon and the phenomenon is called the Photoelectric effect.

If b is comparable to R , the incident photons are scattered by the electron of the atom and the electron itself also gets scattered. This phenomenon is known as Compton effect.

If b is less than R so that the photon is incident in the vicinity of the field of nucleus, the photon is directly converted into an electron-positron pair. This is known as *pair production*.

The above three processes are explained below.

1.5 PHOTOELECTRIC EFFECT

The phenomenon of emission of electrons from the surface of certain substances, mostly metals, when illuminated with visible, ultraviolet light or X-rays, is called photoelectric effect. The emitted electrons are called photoelectrons and the current so produced is called photoelectric current.

Alkali metals like lithium, sodium, potassium, cesium, etc. show photoelectric effect with visible light while metals like zinc, cadmium, magnesium, etc. show photoelectric effect only with ultraviolet radiations or radiations of higher frequencies.

Experimental Demonstration. The experimental demonstration of photoelectric effect can be made with the arrangement shown in Fig. 1.3. It consists of an evacuated quartz tube T with a photosensitive plate C (acts as cathode) and another plate A (acts as anode). The electrodes are connected to a battery B and microammeter (μA) through a rheostat. The tube T with electrodes is also called phototube. A voltmeter V is connected across the tube to measure the potential difference between the electrodes. The microammeter measures the photoelectric current.

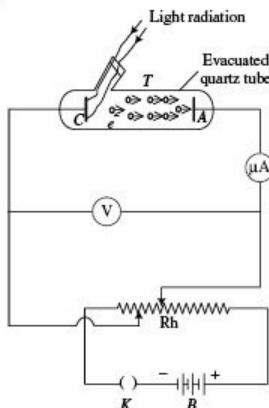


Fig. 1.3

When no light falls on the plate C , no electron is available for conduction of electric current inside the tube (i.e., circuit is not complete), because no photoelectric effect is taking place.

When monochromatic radiations of suitable frequency is made to fall on the plate C , a current is found to be flowing through the circuit. This can be understood as follows. When light of suitable frequency is incident on the plate C , photoelectrons are ejected which are accelerated towards the positively charged plate A . This completes the circuit and a current flows which is measured by the microammeter.

It has been found that the time lag between the incidence of photons and ejection of photoelectrons is less than 10^{-9} s.

(i) Effect of intensity of incident radiation on photoelectric current:

To see the dependence of photoelectric current on the intensity of incident radiation, the potential difference across the two plates is kept constant by adjusting the voltage from battery B with the help of rheostat. The intensity of the incident radiation (having fixed frequency) is varied and the corresponding photoelectric current are noted.

The observed variation of photoelectric current with the intensity of the incident radiation is shown in Fig. 1.4. It is seen that *photoelectric current increases linearly with the intensity of incident radiation*. This implies that *the number of photoelectron emitted is directly proportional to the intensity (or the number of incident photons per unit area) of the incident radiation*.

(ii) Effect of potential on photoelectric current: To find the effect of potential applied on the plate A , the frequency and intensity of the incident radiation is kept constant and photoelectric current of different (anode) potential is measured. The observed variation is depicted in Fig. 1.5. The observation is shown to be repeated for two different intensities I_1 and I_2 .

As seen in Fig. 1.5, initially photoelectric current increases rapidly with increasing positive potential on the anode and as the potential is increased further the photoelectric current saturates. This happens because when light is incident, photoelectrons are emitted. These photoelectrons are accelerated towards positively charged anode. As positive potential on the plate A is increased, more and more photoelectrons are attracted towards it. At a particular anode potential all the photoelectrons being emitted reach the anode and saturation starts. Any further increase in anode potential do not result in an increase in photoelectrons reaching the anode.

When a negative potential is applied on the anode (A) w.r.t. cathode (C) and is gradually increased, photoelectric current is seen (Fig. 1.5) to be decreasing rapidly and finally becomes zero at a certain negative potential. *The minimum negative potential applied to plate A (anode) at which photoelectric current reduces to zero is called cut-off or stopping potential (V_0)*. Below the cut-off potential no photoelectric current flows, no matter how high is the intensity of the incident radiation.

The observed variation can be understood as follows. When a negative potential is applied on plate A , photoelectrons are repelled by it. However, a few photoelectrons having sufficient kinetic energies overcome the repulsive force and pass through the anode constituting a small photoelectric

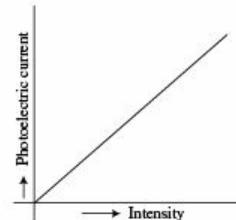


Fig. 1.4

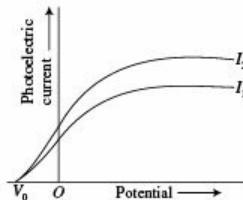


Fig. 1.5

current. As the negative potential is increased, the number of these photoelectrons go on decreasing and at the cut-off potential almost no photoelectron have sufficient kinetic energy to overcome the repulsive force and so photoelectric current reduces to zero.

The stopping potential V_0 is, thus, a measure of the maximum kinetic energy of the photoelectron which is just stopped from reaching the plate C, for a particular incident radiation having a fixed frequency.

If v_{\max} is the maximum velocity of a photoelectron of mass m , then maximum kinetic energy

$$= \frac{1}{2} mv_{\max}^2$$

If V_0 is the stopping potential of this photoelectron, then work done by the repulsive force in stopping this electron = eV_0 ,

where e is charge on an electron.

$$\text{Therefore, we have } \frac{1}{2} mv_{\max}^2 = eV_0$$

The above equation shows that the *maximum velocity of a photoelectron is independent of the intensity of the incident radiation*.

(iii) Effect of frequency of incident radiation: To study the effect of frequency of the incident radiation, radiations of different frequencies but same intensity are made to fall on the plate C. Figure 1.6 shows the observed variation of photoelectric current with the potential applied on plate A for three different frequencies v_1 , v_2 and v_3 ($v_3 > v_2 > v_1$) of the incident radiation. We observe that the *saturation current is the same for the radiation of different frequencies but the value of cut-off potential is different for different frequencies*.

Larger the value of the frequency, greater is the value of cut-off potential. This indicates that when the frequency of the incident radiation (or photon) is higher, the kinetic energies of the emitted electrons are higher and therefore require more repulsive force to be applied to stop them. This is manifested as higher cut-off potential.

A graph plotted between stopping potential and the frequencies of incident radiations is observed to be of the form shown in Fig. 1.7. From the graph, the following two conclusions can be drawn:

- The stopping potential varies linearly with the frequency of the incident radiation.
- Below a certain minimum frequency, called *threshold frequency* (v_0), the stopping potential is zero.

The above observation leads us to conclude that *no emission of photoelectron takes place if the frequency of the incident radiation is less than threshold frequency*. The value of threshold frequency depends upon the nature of the emitting surface and is different for different metals.

Laws of photoelectric emission. Based on the above described experimental observations, the following are the laws of photoelectric emission.

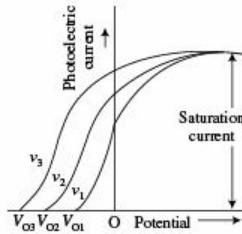


Fig. 1.6

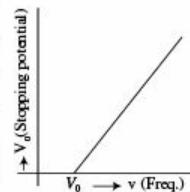


Fig. 1.7

- (i) For a given metal and frequency of radiation, the number of photoelectrons ejected is directly proportional to the intensity of the incident light.
- (ii) For a given metal, there exists a certain minimum frequency of the incident radiation below which no emission of photoelectrons takes place. This frequency is called *threshold frequency*.
- (iii) Above the threshold frequency, the maximum kinetic energy of the emitted photoelectron is independent of the intensity of the incident light and depends only upon the frequency (or wavelength) of the incident light.
- (iv) It is an instantaneous process, i.e., the emission of photoelectrons starts as soon as light falls on the metal surface.

The wave theory of light fails to explain the laws of photoelectric emission. Einstein, using particle nature of light, explained the laws of photoelectric emission with the help of an equation known as Einstein's photoelectric equation, which is discussed below.

1.6 EINSTEIN'S PHOTOELECTRIC EQUATION

Each photon of a radiation of frequency ν carries an energy equal to $h\nu$ (h is Planck's constant). According to Einstein, when a photon of incident radiation collides with an electron, it transfers its energy to the electron. If the frequency of the incident radiation is higher than the threshold frequency (ν_0) for the metal, the energy transferred to the electron by the photon is enough to cause photoelectric emission. A part of the energy (equal to the work function ϕ_0 of the metal) acquired by an electron is spent in overcoming the surface potential barrier and the remaining energy appears as the maximum

kinetic energy $\left(= \frac{1}{2}mv_{\max}^2 \right)$ of the photoelectron. This is represented as Einstein's photoelectric equation, which is given as

$$\frac{1}{2}mv_{\max}^2 = h\nu - \phi_0 \quad \dots(1.10)$$

If ν_0 is the threshold frequency, then $\phi_0 = h\nu_0$

$$\frac{1}{2}mv_{\max}^2 = h\nu - h\nu_0 = h(\nu - \nu_0) \quad \dots(1.11)$$

Using Einstein's photoelectric equation, the laws of photoelectric equation was successfully explained by Einstein for which he was awarded Nobel Prize in Physics in 1913.

Explanation of the laws of photoelectric emission

- (i) **Effect of intensity of incident radiation.** Since each incident photon ($\nu > \nu_0$) ejects one photoelectron from a metal surface, therefore, the number of photoelectrons emitted depends upon the number of photons falling on the metal surface which in turn depend on the intensity of the incident light. If the intensity of the incident light is increased, the number of incident photons increases, which results in an increase in the number of photoelectrons ejected.
- (ii) **Existence of threshold frequency.** If the frequency of the incident radiation is less than the threshold frequency i.e., $\nu < \nu_0$, then the kinetic energy of the electron becomes negative, as evident from Eq. 1.1, which is physically impossible. Hence, photoelectric emission does not take place by the incident radiation having frequency below threshold frequency.

- (iii) **Kinetic energy of photoelectron.** From Eq. 1.1 it is evident that for $v > v_0$, $\frac{1}{2}mv_{\max}^2 \propto v$.

Thus, above the threshold frequency, the maximum kinetic energy of the photoelectrons increases linearly with the frequency of the incident radiation. If the intensity of the incident light is increased, the number of ejected photoelectrons increases but the energy of photoelectrons do not change. Thus, the maximum kinetic energy of the photoelectrons is independent of the intensity of radiation.

- (iv) **Instantaneous process.** Because the photoelectric emission occurs due to elastic collision between a photon and an electron. Practically, there is no time lag ($< 10^{-9}$ s) between the incidence of a photon and emission of an electron.

Example 1.3

Work function of lithium is 2.5 eV. Find the maximum wavelength of light that can cause photoelectric effect in lithium.

Solution. Work function $\phi_0 = hv_0$, or $\phi_0 = \frac{hc}{\lambda_0}$

$$\therefore \lambda_0 = \frac{hc}{\phi_0}$$

Here,

$$\phi_0 = 2.5 \text{ eV} = 2.5 \times 1.6 \times 10^{-19} \text{ J}$$

$$h = 6.63 \times 10^{-34} \text{ Js}, c = 3 \times 10^8 \text{ m/s}$$

\therefore Maximum (threshold) wavelength,

$$\begin{aligned}\lambda_0 &= \frac{6.63 \times 10^{-34} \times 3 \times 10^8}{2.5 \times 1.6 \times 10^{-19}} \text{ m} \\ &= 4.98 \times 10^{-7} \text{ m} \\ &= 4980 \text{ Å}.\end{aligned}$$

Example 1.4

Calculate the kinetic energy of a photoelectron (in eV) emitted on shining light of wavelength 6.2×10^{-6} m on a metal surface. The work function of metal is 0.1 eV ($h = 6.6 \times 10^{-34}$ Js).

Solution.

$$\begin{aligned}\text{Kinetic energy} \quad K_{\max} &= hv - \phi_0 \\ &= \frac{hc}{\lambda} - \phi_0\end{aligned}$$

Here,

$$\lambda = 6.2 \times 10^{-6} \text{ m}$$

$$\phi = 0.1 \text{ eV} = 0.1 \times 1.6 \times 10^{-19} \text{ J}$$

$$\begin{aligned} K_{\max} &= \frac{6.6 \times 10^{-34} \text{ Js} \times 3 \times 10^8 \text{ m/s}}{6.2 \times 10^{-6} \text{ m}} - 0.1 \times 1.6 \times 10^{-19} \text{ J} \\ &= 3.2 \times 10^{-20} \text{ J} - 1.6 \times 10^{-20} \text{ J} \\ &= (3.2 - 1.6) \times 10^{-20} \text{ J} \\ &= 1.6 \times 10^{-20} \text{ J} \\ &= \frac{1.6 \times 10^{-20}}{1.6 \times 10^{-19}} \text{ eV} \\ &= 1 \times 10^{-1} \text{ eV} \\ &= 0.1 \text{ eV} \end{aligned}$$

1.7 PHOTOELECTRIC CELLS

A photoelectric cell is a device for converting light energy into electrical energy.

Its functioning is based on the phenomenon of photoelectric effect. Photoelectric cells are of three types: (i) photoemissive cell, (ii) photovoltaic cell, and (iii) photoconductive cell.

Photoemissive cell: As shown in Fig. 1.8, it consists of an evacuated glass or quartz bulb inside which is a semicylindrical photosensitive metal plate *C* and wire loop *A* which serve as cathode and anode respectively. The anode *A* is in the form of wire loop so that it does not obstruct the path of light. A battery (*B*) is used to apply suitable positive potential on the anode *A*. A microammeter (μA) is used to measure the current.

When light of frequency above the threshold frequency of cathode surface is incident on the cathode, electrons are emitted. Because anode is positively charged, these electrons are attracted towards the anode and move towards the positive terminal of battery, constituting an electric current called photocurrent. The photocurrent is found to be proportional to the intensity of incident light.

Sometimes inert gases like neon or argon are filled so as to increase the photoelectric current due to ionisation. When the potential difference between the electrodes exceed the ionisation potential of the gas, the emitted photoelectrons ionise the gas which increases the magnitude of the current.

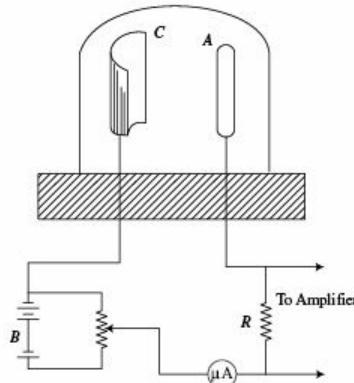


Fig. 1.8

Applications of photoelectric cells. Photoelectric cells have a wide range of applications. A few of these applications are mentioned below:

1. Photocells are used to compare the illuminating powers of two light sources and to find the intensity of illumination of a surface.
2. At the point of transmission of T.V. signals (T.V. studios) photoelectric cells are used to convert light and shade of the object into electric currents for transmission of picture. That is, they are used in *phototelegraphy*.
3. It is used for the reproduction of sound from the soundtrack recorded on the one edge of the cinema films.
4. It is used in *burglar alarm*. A beam of ultraviolet rays is allowed to fall on a photocell across the room. If a burglar enters the room, he unknowingly interrupts the rays falling on the photocell. Thus, the photoelectric current stops for a moment. This switches on another electrical circuit which results in ringing of a bell.
5. It is used to determine the opacity of solids and liquids.
6. It is used in a photographic camera for the automatic adjustment of aperture which in turn controls the amount of light entering the camera.
7. It is used for automatic counting of the number of persons entering a hall, a stadium, etc.
8. It is used for automatic switching of street lights and traffic signals.
9. Photocells are used to control the temperature of furnaces

Example 1.5

For photoelectric effect in sodium, the figure (Fig. 1.9) shows the plot of cut-off voltage versus frequency of incident radiation. Calculate

- (i) the threshold frequency and
- (ii) the work function for sodium.

Solution. (i) The threshold frequency corresponds to the frequency for which the cut-off voltage is zero. Hence, threshold frequency for sodium is 4.5×10^{14} Hz.

(ii) The work function (ϕ_0) is related to the threshold frequency (v_0) by the relation

$$\begin{aligned}\phi_0 &= h v_0 \\ &= 6.6 \times 10^{-34} \times 4.5 \times 10^{14} \text{ J} \\ &= 29.7 \times 10^{-20} \text{ J} \\ &= 2.97 \times 10^{-19} \text{ J} \\ &= \frac{2.97 \times 10^{-19}}{1.6 \times 10^{-19}} \text{ eV} \\ &= 1.856 \text{ eV}\end{aligned}$$

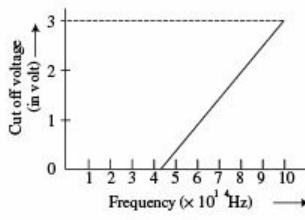


Fig. 1.9

Example 1.6

Light of wavelength 3500 \AA is incident on two metals A and B. Which metal will yield photoelectrons, if their work functions are 4.2 eV and 1.9 eV respectively?

Solution. Here, $\lambda = 3500 \text{ \AA} = 3500 \times 10^{-10} \text{ m}$

$$\begin{aligned}\text{Energy of incident photons} &= h\nu = \frac{hc}{\lambda} \\ &= \frac{6.6 \times 10^{-34} \times 3 \times 10^8}{3500 \times 10^{-10} \times 1.6 \times 10^{-19}} \text{ eV} \\ &= \frac{6.6 \times 3}{3.5 \times 1.6} \text{ eV} \\ &= 3.536 \text{ eV.}\end{aligned}$$

The work function of metal A being higher than the energy of incident photon it will not yield any photoelectron; only metal A will yield photoelectrons.

Example 1.7

Find the frequency of light, which ejects electrons from a metal surface fully stopped by a retarding potential of 3 V . The photoelectric effect begins in this metal at a frequency of $6 \times 10^{14} \text{ s}^{-1}$. Find the work function of this metal.

Solution. Here, $V_0 = 3 \text{ volt}$, $V_0 = 6 \times 10^{14} \text{ s}^{-1}$

$$v = ? \text{ and } \phi_0 = ?$$

$$h\nu = h\nu_0 + \frac{1}{2} mv_{\max}^2$$

Also,

$$eV_0 = \frac{1}{2} mv_{\max}^2$$

∴

$$h\nu = h\nu_0 + eV_0$$

or,

$$v = v_0 + \frac{eV_0}{h}$$

$$= 6 \times 10^{14} \text{ s}^{-1} + \frac{1.6 \times 10^{-19} \times 3}{6.6 \times 10^{-34}} \text{ s}^{-1}$$

$$= 6 \times 10^{14} \text{ s}^{-1} + 0.727 \times 10^{15} \text{ s}^{-1}$$

$$= 6 \times 10^{14} \text{ s}^{-1} + 7.27 \times 10^{14} \text{ s}^{-1}$$

$$= (6.00 + 7.27) \times 10^{14} \text{ s}^{-1}$$

$$= 13.27 \times 10^{14} \text{ s}^{-1}$$

$$\begin{aligned}
 \text{Work function} \quad \phi_0 &= h\nu_0 \\
 &= 6.6 \times 10^{-34} \text{ Js} \times 6 \times 10^{14} \text{ s}^{-1} \\
 &= 3.96 \times 10^{-19} \text{ J} \\
 &= \frac{3.96 \times 10^{-19}}{1.6 \times 10^{-19}} \text{ eV} \\
 &= 2.475 \text{ eV.}
 \end{aligned}$$

Example 1.8

Monochromatic radiation of wavelength 640.2 nm ($1 \text{ nm} = 10^{-9} \text{ m}$) from a neon lamp irradiates a photosensitive material made of cesium on tungsten. The stopping voltage is measured to be 0.54 V. The source is replaced by an iron source and its 427.2 nm light irradiates the same photocell. Predict the new stopping voltage.

Solution.

$$\begin{aligned}
 \lambda &= 640.2 \text{ nm} = 640.2 \times 10^{-9} \text{ m} \\
 V_0 &= 0.54 \text{ volt} \\
 \text{Einstein photoelectric eq.} \quad h\nu &= \phi_0 + \frac{1}{2} mv_{\max}^2 \\
 \therefore \text{Work function} \quad \phi_0 &= h\nu - \frac{1}{2} mv_{\max}^2 \\
 \text{As} \quad \frac{1}{2} mv_{\max}^2 &= eV_0 \\
 \phi_0 &= h\nu - eV_0 \\
 &= \frac{hc}{\lambda} - eV_0 \\
 &= \frac{6.62 \times 10^{-34} \times 3 \times 10^8}{640.2 \times 10^{-9}} - 1.6 \times 10^{-19} \times 0.54 \\
 &= 3.102 \times 10^{-19} - 0.864 \times 10^{-19} \text{ J} \\
 &= (3.102 - 0.864) \times 10^{-19} \text{ J} \\
 &= 2.238 \times 10^{-19} \text{ J} \\
 \text{With iron source} \quad \lambda &= 427.2 \text{ nm} = 427.2 \times 10^{-9} \text{ m} \\
 V_0 &=? \\
 \text{From photoelectric eq., we have} \quad eV_0 &= h\nu - \phi_0 \\
 V_0 &= \frac{1}{e} \left(\frac{hc}{\lambda} - \phi_0 \right)
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{1.6 \times 10^{-19}} \left(\frac{6.62 \times 10^{-34} \times 3 \times 10^8}{427.2 \times 10^{-19}} - 2.238 \times 10^{-19} \right) V \\
 &= \frac{1}{1.6 \times 10^{-19}} (4.649 \times 10^{-19} - 2.238 \times 10^{-19}) V \\
 &= \frac{2.411 \times 10^{-19}}{1.6 \times 10^{-19}} = 1.507 \text{ V.}
 \end{aligned}$$

Example 1.9

When ultraviolet light of wavelength 227.1 nm irradiates a photocell made of molybdenum metal from a 100 W mercury source, the stopping potential is found to be 1.3 V. How would the photocell respond to a high intensity ($\approx 10^6 \text{ W m}^{-2}$) red light of wavelength 6328 Å produced by a He-Ne laser?

Solution.

Here,

$$\lambda = 227.1 \text{ nm} = 227.1 \times 10^{-9} \text{ m}$$

$$V_0 = 1.3 \text{ V}$$

\therefore Work function of molybdenum metal,

$$\begin{aligned}
 \phi_0 &= h\nu - eV_0 \\
 &= \frac{hc}{\lambda} - eV_0 \\
 \phi_0 &= \frac{6.62 \times 10^{-34} \times 3 \times 10^8}{227.1 \times 10^{-9}} - 1.6 \times 10^{-19} \times 1.3 \text{ J} \\
 &= 8.745 \times 10^{-9} - 2.08 \times 10^{-19} \text{ J} \\
 &= (8.745 - 2.080) \times 10^{-19} \text{ J} \\
 &= 6.665 \times 10^{-19} \text{ J} \\
 &= \frac{6.665 \times 10^{-19}}{1.6 \times 10^{-19}} \text{ eV} \\
 &= 4.166 \text{ eV}
 \end{aligned}$$

\therefore Threshold wavelength

$$\lambda_0 = \frac{hc}{\phi_0}$$

$$\left(\begin{array}{l} h\nu_0 = \phi_0 \\ \text{or } \frac{hc}{\lambda_0} = \phi_0 \end{array} \right)$$

$$\begin{aligned}
 &= \frac{6.62 \times 10^{-34} \times 3 \times 10^8}{6.665 \times 10^{-19}} \text{ m} \\
 &= 2.98 \times 10^{-7} \text{ m} \\
 &= 2980 \text{ Å}
 \end{aligned}$$

Since wavelength 6328 Å of incident red light from He-Ne laser is greater than the threshold wavelength 2980 Å of molybdenum metal, no photoemission will take place in the photocell.

Example 1.10

Work function of sodium is 2.35 eV. What is the maximum wavelength of light that will cause photoelectrons to be emitted from the metal? What will be the maximum energy of the photoelectrons, if radiation of 1000 Å falls on the metal surface ($\hbar = 6.625 \times 10^{-34}$ Js).

Solution.

Here,

$$\begin{aligned}\phi_0 &= 2.35 \text{ eV} \\ &= 2.35 \times 1.6 \times 10^{-19} \text{ J} \\ \lambda &= 1000 \text{ Å} \\ &= 10^{-7} \text{ m}\end{aligned}$$

We know

$$\phi_0 = h\nu_0 = \frac{hc}{\lambda_0} \quad \left(\begin{array}{l} \nu_0 = \text{threshold freq.} \\ \lambda_0 = \text{threshold wavelength} \end{array} \right)$$

$$\begin{aligned}\therefore \text{Maximum wavelength} \quad \lambda_0 &= \frac{hc}{\phi_0} \\ &= \frac{6.625 \times 10^{-34} \times 3 \times 10^8}{2.35 \times 1.6 \times 10^{-19}} \text{ m} \\ &= 5.286 \times 10^{-7} \text{ m} \\ &= 5286 \text{ Å}\end{aligned}$$

Maximum energy of photoelectrons,

$$\begin{aligned}\frac{1}{2}mv_{\max}^2 &= h\nu - \phi_0 \\ \text{or,} \quad \frac{1}{2}mv_{\max}^2 &= \frac{hc}{\lambda} - \phi_0 \\ \therefore \text{Max. energy} &= \frac{6.625 \times 10^{-34} \times 3 \times 10^8}{10^{-7}} - 2.35 \times 1.6 \times 10^{-19} \text{ J} \\ &= 19.875 \times 10^{-19} - 3.76 \times 10^{-19} \text{ J} \\ &= (19.875 - 3.760) \times 10^{-19} \text{ J} \\ &= 16.115 \times 10^{-19} \text{ J} \\ &= \frac{16.115 \times 10^{-19}}{1.6 \times 10^{-19}} \text{ eV} \\ &= 10.07 \text{ eV}\end{aligned}$$

Example 1.11

When light of wavelength 400 nm is incident on the cathode of a photocell, the stopping potential recorded is 6 V. If the wavelength of the incident light increases to 600 nm, calculate the new stopping potential.

Solution.

Here

$$\lambda_1 = 400 \text{ nm} = 400 \times 10^{-9} \text{ m}, V_{S1} = 6 \text{ V}$$

$$\lambda_2 = 600 \text{ nm} = 600 \times 10^{-9} \text{ m}, V_{S2} = ?$$

$$V_{S1} - V_{S2} = \frac{hc}{e} \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} \right)$$

or

$$\begin{aligned} V_{S2} &= V_{S1} - \left[\frac{hc}{e} \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} \right) \right] \\ &= 6 \text{ V} - \left[\frac{6.6 \times 10^{-34} \times 3 \times 10^8}{1.6 \times 10^{-19}} \left(\frac{1}{400 \times 10^{-9}} - \frac{1}{600 \times 10^{-9}} \right) \right] \\ &= 6 \text{ V} - 1.03 \text{ V} \\ &= 4.97 \text{ V}. \end{aligned}$$

Example 1.12

Light of intensity 10^{-5} W m^{-2} falls on a sodium photocell of surface area 2 cm^2 . Assuming that the top 5 layers of sodium absorb the incident energy, estimate the time required for photoelectric emission in the wave picture of radiation. The work function for the metal is given to be about 2 eV. What is the implication of your answer?

Solution. Sodium has one conduction electron available per atom.

$$\text{Effective atomic area} \approx 10^{-20} \text{ m}^2$$

Number of conduction electrons in 5 layers

$$\begin{aligned} &= \frac{\text{Area of a layer}}{\text{Effective atomic area}} \\ &= \frac{5 \times 2 \times 10^{-4}}{10^{-20} \text{ m}^2} = 10^{17} \end{aligned}$$

$$\text{Intensity} = \frac{\text{Energy}}{\text{Area} \times \text{Time}} = \frac{\text{Power}}{\text{Area}}$$

$$\therefore \text{Incident power} = \text{Intensity} \times \text{area}$$

$$= 10^{-5} \text{ W m}^{-2} \times 2 \times 10^{-4} \text{ m}^2 = 2 \times 10^{-9} \text{ W}$$

In the wave picture, the incident light energy is uniformly absorbed by all the electrons.
 \therefore Energy absorbed per second per electron

$$= \frac{\text{Incident power}}{\text{Number of electrons}} = \frac{2 \times 10^{-9}}{10^{17}}$$

$$= 2 \times 10^{-26} \text{ W}$$

Time required for photoelectric emission

$$= \frac{\text{Energy required per electron}}{\text{Energy absorbed per second per electron}}$$

$$= \frac{2 \text{ eV}}{2 \times 10^{-26} \text{ W}} = \frac{2 \times 1.6 \times 10^{-19} \text{ J}}{2 \times 10^{-26} \text{ Js}^{-1}} = 1.6 \times 10^7 \text{ s} \cong 0.5 \text{ year}$$

The implication is that the wave picture of radiation is in gross disagreement with experimental observation as photoelectric emission is observed to be a nearly spontaneous ($= 10^{-9} \text{ s}$) phenomenon.

1.8 COMPTON EFFECT

In 1921, Prof. A.H. Compton discovered that when a monochromatic beam of high frequency radiation (e.g., X -rays and γ -rays) is scattered by a substance, the scattered radiation contains the radiation of greater wavelength (or lower frequency) along with the radiation of the incident wavelength (or frequency). The radiation of unchanged wavelength in the scattered light are called *unmodified radiations* while the radiations of greater wavelength are called *modified radiations*. The phenomenon is called *Compton effect*.

Compton's Experimental Arrangement

The concept of a photon as a concentrated packet of energy travelling with the speed of light was confirmed by Compton in 1923. He allowed X -rays of monochromatic wavelength λ to fall on a graphite block and measured the intensity of scattered X -rays at different angles. The scattered X -rays are found to have two wavelengths—the original wavelength λ and another wavelength λ' , λ' being greater than λ . The schematic experimental arrangement is shown in Fig. 1.10 (a).

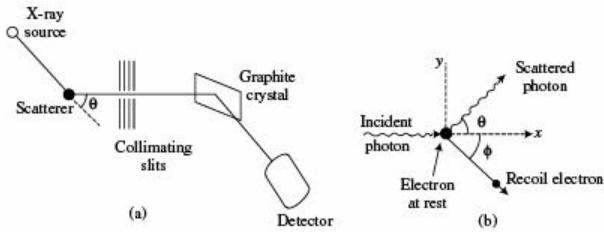


Fig. 1.10 (a) Experimental arrangement and (b) Scattering process in Compton effect.

Explanation

The effect was explained on the basis of quantum theory of light. The wave picture predicts scattered radiation having the same wavelength though less energetic than the incident radiation. However, the quantum theory, i.e., the photon concept leads to exact result, as discussed below.

Compton analyzed the scattering process as a particle like collision between a single photon of energy $h\nu$ and an electron which is assumed to be at rest. During collision, the incident photon transfers some of its energy to the electron which is almost free. Hence, the scattered photon must have a lower energy $h\nu'$ ($\nu' < \nu$). This implies a higher wavelength λ' . The scattering process is illustrated in Fig. 1.10 (b).

Resolving the momenta of the incident photon, scattered photon and electron into x and y components and applying the law of conservation of momentum, we have

$$\frac{h\nu}{c} = \frac{h\nu'}{c} \cos \theta + p_e \cos \phi \quad \dots(1.12)$$

$$0 = \frac{h\nu'}{c} \sin \theta - p_e \sin \phi \quad \dots(1.13)$$

where p_e is the momentum of the electron after scattering. Conservation of energy gives

$$h\nu = h\nu' + k_e$$

or $k_e = h(\nu - \nu') \quad \dots(1.14)$

where k_e is the kinetic energy of scattered electron. Eqs. (1.12) and (1.13) can be written as

$$p_e c \cos \phi = h\nu - h\nu' \cos \theta \quad \dots(1.15)$$

$$p_e c \sin \phi = h\nu' \sin \theta \quad \dots(1.16)$$

Squaring and adding, we get

$$p_e^2 c^2 = h^2 \nu^2 + h^2 \nu'^2 - 2h^2 \nu \nu' \cos \theta \quad \dots(1.17)$$

Denoting the rest mass of electron by m_0 , from Eq. (1.14), we have

$$k_e + m_0 c^2 = h(\nu - \nu') + m_0 c^2 \quad \dots(1.18)$$

From the theory of relativity, the energy of the recoiling electron is

$$E_e = (c^2 p_e^2 + m_0^2 c^4)^{1/2} \quad \dots(1.19)$$

Since left-hand side of Eq. (1.18) is the total energy E_e of the recoiling electron.

\therefore From Eq. (1.18) and using Eq. (1.19), we have

$$(c^2 p_e^2 + m_0^2 c^4)^{1/2} = h(\nu - \nu') + m_0 c^2 \quad \dots(1.20)$$

Squaring Eq. (1.20) and using Eq. (1.17), we get

$$h^2 \nu^2 + h^2 \nu'^2 - 2h^2 \nu \nu' \cos \theta = h^2 (\nu - \nu')^2 + 2h(\nu - \nu') m_0 c^2$$

or $h^2 \nu \nu' \cos \theta - h^2 \nu \nu' + h(\nu - \nu') m_0 c^2 = 0 \quad \dots(1.21)$

Since $\nu = c/\lambda$ and $\nu' = c/\lambda'$, Eq. (1.21) reduces to

$$\frac{h^2 c^2}{\lambda \lambda'} \cos \theta - \frac{h^2 c^2}{\lambda \lambda'} + h(c/\lambda - c/\lambda') m_0 c^2 = 0$$

Multiplying both sides by $\lambda\lambda'$ and rearranging, we have

$$\begin{aligned} h(1 - \cos \theta) &= c(\lambda' - \lambda) m_0 \\ (\lambda' - \lambda) &= \frac{h}{m_o c} (1 - \cos \theta) \end{aligned} \quad \dots(1.22)$$

$$\text{or, } \Delta\lambda = \frac{h}{m_o c} (1 - \cos \theta) \quad \dots(1.22\text{a})$$

This is the expression for the *Compton shift*.

Since the right-hand side of Eq. (1.22) is always positive, λ' is always greater than λ . The change in wavelength depends upon angle of scattering. The quantity $(h/m_o c)$ is referred as *Compton wavelength* of the electron. Substituting the values,

$$\frac{h}{m_o c} = \frac{6.6 \times 10^{-34}}{9 \times 10^{-31} \times 3 \times 10^8} \text{ m} = 0.02426 \text{ Å}^\circ.$$

The maximum value of $(1 - \cos \theta)$ (for $\theta = 180^\circ$) is 2. Therefore, the maximum possible change in wavelength is 0.04852 Å or 0.05 Å approximately. It follows from Eq. (1.22) that the Compton effect can most readily be detected for radiation of wavelength not greater than a few angstrom units. For example, for $\lambda = 5 \text{ Å}$, the maximum change in wavelength is 1% while for $\lambda = 1 \text{ Å}$ the effect is 5%. For visible light (taking $\lambda \approx 5000 \text{ Å}$), the maximum wavelength change ($\approx 0.05 \text{ Å}$) is only about 0.001% of the incident wavelength which is undetectable. Hence, *Compton effect cannot be observed for visible light rays*.

Further, it is assumed that the electron is so loosely bound to the atom that it can be regarded as free, which is not true and the electron is tightly bound to the atom and the atom as a whole recoils due to Compton scattering. Therefore, in conservation relations instead of m_0 , the rest mass of M_0 of the atom should be used. For an aluminium target $M_0 \approx 27 \frac{m_H}{m_o} \times m_o = 27 \times 1840 m_o$, where m_H is the mass of the hydrogen atom. Then

$$\frac{h}{m_0 c} = \frac{2.246 \times 10^{-12}}{27 \times 1840} \approx 4.9 \times 10^{-17} \text{ m} \approx 4.9 \times 10^{-7} \text{ Å}$$

Thus, for an incident wavelength of the order of a few Å, when electron is bound to the atom, the change in wavelength is negligible for all θ . This explains the presence of unmodified radiation.

Direction of Recoil Electron

Dividing Eq.(1.16) by Eq. (1.15), we get

$$\begin{aligned} \tan \phi &= \frac{hv' \sin \theta}{hv - hv' \cos \theta} = \frac{\frac{c}{\lambda'} \sin \theta}{\frac{c}{\lambda} - \frac{c}{\lambda'} \cos \theta} \\ \tan \phi &= \frac{\lambda \sin \theta}{(\lambda' - \lambda \cos \theta)} \end{aligned} \quad \dots(1.23)$$

Kinetic Energy of the Recoil Electron

The kinetic energy of the recoil electron is given by

$$k_e = (m - m_o)c^2 \quad \dots(1.24)$$

Also,

$$k_e = h\nu - h\nu' \quad \dots(1.25)$$

or

$$k_e = \frac{hc}{\lambda} - \frac{hc}{\lambda'} = hc \left(\frac{\lambda' - \lambda}{\lambda \lambda'} \right) \quad \dots(1.26)$$

From Eq. (1.22), we have

$$\lambda' = \lambda + \frac{h}{m_o c} (1 - \cos \theta)$$

or

$$\frac{c}{\nu'} = \frac{c}{\nu} + \frac{h}{m_o c} (1 - \cos \theta)$$

$$\frac{1}{\nu'} = \frac{1}{\nu} \left[1 + \frac{h\nu}{m_o c^2} (1 - \cos \theta) \right]$$

$$\nu' = \frac{\nu}{1 + \frac{h\nu}{m_o c^2} (1 - \cos \theta)} = \frac{\nu}{1 + x (1 - \cos \theta)} \quad \dots(1.27)$$

where,

$$x = \frac{h\nu}{m_o c^2}$$

Substituting value of ν' from Eq. (1.27) in Eq. (1.25), the kinetic energy of the recoil electron is

$$\begin{aligned} k_e &= h\nu - h \left[\frac{\nu}{1 + x (1 - \cos \theta)} \right] \\ &= h\nu \left[1 - \frac{1}{1 + x (1 - \cos \theta)} \right] \\ k_e &= h\nu \left[\frac{x (1 - \cos \theta)}{1 + x (1 - \cos \theta)} \right] \end{aligned} \quad \dots(1.28)$$

where,

$$x = \frac{h\nu}{m_o c^2}$$

Variation in Compton shift as θ varies from 0° to 180°

The change in wavelength $\Delta\lambda (= \lambda' - \lambda)$ changes from zero to twice the Compton wavelength as θ changes from 0° to 180° . We have

$$\lambda' - \lambda = \frac{h}{m_a c} (1 - \cos \theta)$$

$$\text{For } \theta = 0^\circ, \lambda' - \lambda = \Delta\lambda = \frac{h}{m \cdot c} (0) = 0$$

$$\text{For } \theta = 90^\circ, \lambda' - \lambda = \Delta\lambda = \frac{h}{m_e c} (1 - 0) = \frac{h}{m_e c}$$

$$\text{For } \theta = 180^\circ, \lambda' - \lambda = \Delta \lambda = \frac{h}{m_o c} (1 - (-1)) = \frac{2h}{m_o c}.$$

This is illustrated in Fig. 1.11.

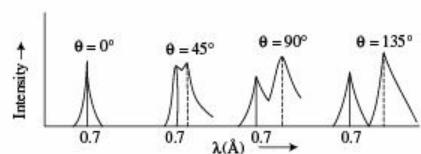


Fig. 1.11 Variation of λ' with θ in Compton experiment. Solid vertical line corresponds to λ and dotted line that to λ' .

Example 1.13

X-rays of wavelength 0.15 nm are scattered from a block of carbon. Calculate wavelength of X-rays scattered at

Solution.

Compton wavelength λ' of scattered wave is given by

$$\lambda' = \lambda + \frac{h}{m_e c} (1 - \cos \theta)$$

- (i) For $\theta = 0^\circ$, $\cos \theta = 1 \quad \therefore \quad \lambda' = \lambda = 0.15 \text{ nm}$

(ii) For $\theta = 90^\circ$, $\cos 90^\circ = 0$

$$\begin{aligned}\lambda' &= \lambda + \frac{2h}{m_e c} \\ &= 0.15 + \frac{6.626 \times 10^{-34} \text{ Js}}{(9.1 \times 10^{-31} \text{ kg})(3 \times 10^8 \text{ ms}^{-1})} \\ &= 0.15 + 2.4 \times 10^{-12} \text{ m} \\ &= 0.1524 \text{ nm}\end{aligned}$$

- (iii) For $\theta = 180^\circ$, $\cos 180^\circ = -1$

$$\begin{aligned}\therefore \lambda' &= \lambda + \frac{2h}{m_o c} \\ &= 0.15 + 2 \times 2.4 \times 10^{-12} \\ &= 0.155 \text{ nm}\end{aligned}$$

Example 1.14

X-rays with wavelength $\lambda = 1.00 \text{ \AA}$ are scattered from a carbon block. The scattered radiations are viewed at right angles to the direction of incident beam. Calculate

- (i) the Compton shift $\Delta\lambda$ in \AA ,
- (ii) wavelength of scattered radiation, and
- (iii) the energy imparted to the recoiling electron in Joules.

Given: m_o (rest mass of e^-) = $9.1 \times 10^{-31} \text{ kg}$, $c = 3 \times 10^8 \text{ m/s}$, $h = 6.6 \times 10^{-30} \text{ Js}$

Solution.

- (i) Compton shift $\Delta\lambda$ is given by

$$\begin{aligned}\Delta\lambda &= \lambda' - \lambda = \frac{h}{m_o c} (1 - \cos \theta) \\ &= \frac{6.6 \times 10^{-34}}{9.1 \times 10^{-31} \times 3 \times 10^8} (1 - \cos 90^\circ) \\ &= 2.417 \times 10^{-12} (1 - 0) \\ &= 0.02417 \text{ \AA}\end{aligned}$$

- (ii) Wavelength of scattered radiation

$$\begin{aligned}\lambda' &= \lambda + \Delta\lambda = (1 + 0.02417) \text{ \AA} \\ &= 1.02417 \text{ \AA}\end{aligned}$$

- (iii) Energy of the recoiled electron

$$k_e = h\nu \left[\frac{x(1 - \cos \theta)}{1 + x(1 - \cos \theta)} \right] s$$

where

$$x = \frac{h\nu}{c^2}$$

$$\nu = \frac{c}{\lambda} = \frac{3 \times 10^8}{1 \times 10^{-10}} = 3 \times 10^{18} \text{ s}^{-1}$$

$$x = \frac{6.6 \times 10^{-34} \times 3 \times 10^{18}}{9.1 \times 10^{-31} \times (3 \times 10^8)^2} = 0.024$$

$$\therefore k_e = 6.6 \times 10^{-34} \times 3 \times 10^{18} \times \left[\frac{0.024 \times 1}{1 + (0.024 \times 1)} \right]$$

$$= 4.64 \times 10^{-14} \text{ Joules}$$

\Rightarrow Energy imparted to the recoiled electron is 4.64×10^{-14} Joules.

Example 1.15

In Compton scattering θ and $\Delta\lambda$, i.e., angle of scattering and Compton shift are related as

$$\Delta\lambda = \frac{h}{m_e c} (1 - \cos \theta)$$

if incident photon has wavelength 2.0×10^{-10} m and $\phi = 90^\circ$. Calculate the angle at which the recoil electron appears.

Solution.

$$\tan \theta = \frac{\lambda \sin \theta}{\lambda' - \lambda \cos \theta}$$

$$= \frac{2 \times 10^{-10} \sin 90^\circ}{2.204 \times 10^{-10} - 2 \times 10^{-10} \cos 90^\circ} = \frac{2}{2.024} = 0.9980$$

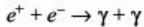
$$\theta = \tan^{-1}(0.9980) = 44^\circ 57'$$

1.9 PAIR PRODUCTION AND ANNIHILATION

When γ -ray photon passes close to an atomic nucleus and it possess energy greater than 1.02 MeV, it gets converted into electron-positron pair. This process is termed *pair production* and in this process basically electromagnetic energy gets converted into matter. This phenomenon was observed by Anderson in 1932.

This process is possible only when the incident photon energy exceeds the rest mass energies of both electron and positron (i.e., photon energy > 0.51 eV + 0.51 eV, rest mass energy of both electron and positron is 0.51 eV). The corresponding maximum photon wavelength is 1.2 pm. This phenomenon occurs mainly in the nuclear field. Both charge and momentum are conserved in this process, however, the presence of nucleus is necessary for the conservation of momentum which carries away most of the photon momentum for the process to occur. Because of its relatively enormous mass, the nucleus absorbs only a negligible fraction of photon energy.

The inverse of pair production occurs when a positron is near an electron and the two come together under the influence of their opposite electric charges. Both the particles vanish simultaneously with lost mass converting into energy of the two photons. The process, termed *pair annihilation*, can be represented as follows.



The directions of the photons are such that both energy and linear momentum are conserved and no nucleus or other particle is needed for pair annihilation.

1.10 DUAL NATURE OF RADIATION: DE-BROGLIE WAVES-MATTER WAVES

Light waves show interference, diffraction and polarisation. These phenomena can be explained only on the basis of wave nature of light. On the other hand, photoelectric effect, Compton effect, Raman effect, etc. can be explained only on the basis of particle nature of electromagnetic radiations.

Thus, in some physical phenomena, electromagnetic radiations exhibit wave nature while in some other phenomena it exhibits particle nature. In other words, electromagnetic radiations exhibit *dual nature*.

In 1924, Louis de Broglie suggested that matter, like radiations, have dual nature, i.e., molecules, atoms, electrons, protons, etc. exhibit wave-like behaviour under certain conditions. He put forward the following arguments to defend his duality principle:

1. The entire energy in the universe is in the form of matter and electromagnetic radiation.
2. *The nature loves symmetry. As the radiation has got dual nature, the matter should also possess dual nature.*

Thus, according to de-Broglie, a wave is associated with every moving material particle. These waves, associated with moving material particles, are called *de-Broglie waves or matter waves*. Wavelength is an essential feature of any wave. The expression for wavelength of de Broglie waves is derived below.

According to quantum theory of radiation, the energy (E) of a photon of radiation of frequency ν is given by

$$E = h\nu \quad \dots(1.29)$$

where, h is Planck's constant.

The relativistic energy of a particle of rest mass m_o is given by

$$E = \sqrt{m_o^2 c^4 + p^2 c^2}$$

For a photon, rest mass $m_o = 0$

$$E = pc \quad \dots(1.30)$$

From Eqs. (1.29) and (1.30), we have

$$pc = h\nu$$

or

$$p = \frac{h\nu}{c}$$

or,

$$p = \frac{h\nu}{\lambda}$$

or,

$$p = \frac{h}{\lambda}$$

$$\therefore \text{de-Broglie wavelength, } \lambda = \frac{h}{p} \quad \dots(1.31)$$

de-Broglie asserted that this is a general formula applicable to any moving particle having momentum. For a particle having mass m and moving with velocity v ,

$$p = mv$$

$$\therefore \lambda = \frac{h}{mv} \quad \dots(1.32)$$

This is known as *de-Broglie equation*.

From Eq. (1.32), the following conclusions can be drawn about the matter waves.

- (i) Lighter is the material particle, longer is its associated de-Broglie wavelength.
- (ii) The faster the particle moves, smaller is its de Broglie wavelength.
- (iii) de-Broglie wavelength of a particle is not dependent on its charge.

It should be remembered that matter waves are not electromagnetic in nature. Electromagnetic waves are produced by accelerating charges only.

If the velocity of the particle (v) is comparable to the velocity of light (c), then mass of the particle is given by

$$m = \frac{m_o}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Therefore, de-Broglie relation in such a case is given by

$$\begin{aligned} \lambda &= \frac{h}{\frac{m_o}{\sqrt{1 - \frac{v^2}{c^2}}} v} \\ \text{or, } \lambda &= \frac{h \sqrt{1 - \frac{v^2}{c^2}}}{m_o v} \end{aligned} \quad \dots(1.33)$$

1.11 DE-BROGLIE WAVELENGTH OF ELECTRON

Consider an electron of mass m and charge e . Let v be the velocity acquired by it when it is accelerated through a potential difference of V volt.

$$\text{Kinetic energy gained by the electron} = \frac{1}{2} mv^2$$

$$\text{Work done on the electron} = eV$$

Applying law of conservation of energy, we have

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

$$\text{or, } v = \sqrt{\frac{2eV}{m}}$$

de-Broglie wavelength of the electron is given by

$$\lambda = \frac{h}{mv}$$

Substituting for v from Eq. (1.34), we get

$$\lambda = \frac{h}{m\sqrt{\frac{2eV}{m}}} = \frac{h}{\sqrt{2meV}} \quad \dots(1.35)$$

Putting $h = 6.62 \times 10^{-34}$ Js, $m = 9.1 \times 10^{-31}$ kg, $e = 1.6 \times 10^{-19}$ C, we get

$$\begin{aligned} \lambda &= \frac{6.62 \times 10^{-34}}{\sqrt{2 \times 9.1 \times 10^{-31} \times 1.6 \times 10^{-19} \times V}} \\ &= \frac{12.27 \times 10^{-34}}{\sqrt{V}} \text{ m} \\ &= \frac{12.27}{\sqrt{V}} \text{ Å} \end{aligned}$$

Suppose, $V = 100$ volt, then de-Broglie wavelength

$$\lambda = \frac{12.27}{\sqrt{100}} \text{ Å} = 1.227 \text{ Å}$$

This is of the order of the wavelength of X-rays ($\approx 1 \text{ Å}$). However, the two wavelengths are quite different in nature.

1.12 DAVISSON AND GERMER EXPERIMENT—VERIFICATION OF MATTER WAVES

The wave nature of slowly moving electrons was established experimentally by Davisson and Germer in 1927. This was confirmed by G.P. Thomson in 1928.

Experimental arrangement used by Davisson and Germer is shown in Fig. 1.12. It consists of a filament F of tungsten coated with barium oxide, which on heating emits electrons. It acts as cathode. A is a cylinder with fine hole along its axis. It is kept at a positive potential w.r.t. cathode and is called anode. The cathode and anode form an electron gun, by which a fine beam of electrons can be obtained under different accelerating potentials. N is a nickel crystal cut along cubical diagonal. D is an electron detector which can be rotated on a circular scale. A sensitive galvanometer (not shown in the figure) is connected to it to record the current.

A fine beam of accelerated electrons is made to fall normally on the nickel crystal. The incident electrons are scattered in different directions by the atoms of the crystal. The intensity of the

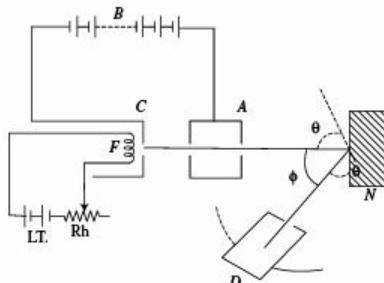


Fig. 1.12

scattered electrons in a direction is found by the use of detector by rotating the detector on a circular scale at different positions, the intensity of the scattered beam is measured for different values of latitude angle (or angle of scattering), the angle between the incident and the scattered electron beam. Graphs are plotted between ϕ and the intensity of scattered beam, shown in Fig. 1.13.

From these curves, following observations can be made.

- Intensity of scattered electron depends upon the angle of scattering.
- A bump (or a kink) always occurs in the curve at $\phi = 50^\circ$, the angle which the scattered beam makes with the incident beam.
- The size of bump is maximum when accelerating potential is 54 V, decreasing both for lower and higher values of accelerating potential.

The appearance of bump in a particular direction is due to constructive interference of electrons scattered from different layers of atoms in the crystal. This established the wave nature of electron. From Fig. 1.20, we note that for the latitude angle $\phi = 50^\circ$, the angle of glancing θ (angle between the scattered beam of electron with the plane of atoms of the crystal) for the electron beam is given by

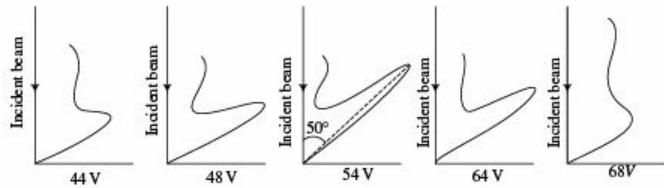


Fig. 1.13

$$\theta + \phi + \theta = 180^\circ$$

$$\text{or, } \theta = \frac{1}{2}(180^\circ - \phi) = \frac{1}{2}(180 - 50) = 65^\circ$$

For Ni crystal, d (interatomic separation) = 0.91 Å

According to Bragg's law of diffraction, $2d \sin \theta = n\lambda$

For first order diffraction, $n = 1$

$$\text{As } 2d \sin \theta = \lambda$$

$$\therefore \lambda = 2 \times 0.91 \times \sin 65^\circ = 1.65 \text{ Å}$$

According to de-Broglie hypothesis, the wavelength of the wave associated with electron is given by

$$\begin{aligned} \lambda &= \frac{h}{mv} = \frac{12.27}{\sqrt{V}} \text{ Å} \\ &= \frac{12.27}{\sqrt{54}} \text{ Å} = 1.66 \text{ Å} \end{aligned}$$

Thus, the experimentally measured wavelength is very close to that calculated from de-Broglie hypothesis. This establishes the existence of de-Broglie waves or matter waves.

1.13 WAVE PACKET: GROUP VELOCITY AND PHASE (OR WAVE) VELOCITY

According to de-Broglie hypothesis a moving material particle has a wavelength $\lambda = \frac{h}{mv}$ associated with it, where m is the mass of the particle and v its velocity. If E is the energy of the particle, then the frequency ν is given by

$$\nu = \frac{E}{h} = \frac{mc^2}{h} \quad (\therefore E = mc^2 - \text{Einstein's mass energy relation})$$

Also, frequency

$$\nu = \frac{v}{\lambda} \quad \text{or} \quad \nu = v\lambda$$

$$\therefore \nu = \frac{mc^2}{h} \frac{h}{mv} = \frac{c^2}{v} = c \times \frac{c}{v}$$

As

$$\frac{c}{v} > 1 \Rightarrow \nu > c$$

That is, the wave associated with the particle would travel faster than the particle itself for which the speed of light is the ultimate. This implies that the material particle cannot be equivalent to a single wave train. The paradox was resolved by Schrodinger by postulating that *a moving material particle is equivalent to a wave packet rather than a single wave*.

A wave packet consists of a group of waves, each of slightly different velocity and wavelength, with such phases and amplitudes that they interfere, constructively over only a small region of space where the particle can be located, outside of which they produce destructive interference so that the amplitude reduces to zero. Such a wave packet is shown in Fig. 1.14. Such a packet moves with a velocity, called the group velocity. The individual wave forming the packet possess an average velocity, called the *phase (or wave) velocity*.

Suppose, we are looking at waves of the sea in front of us as they are propagating from the left-to right-hand side. In order to measure the velocity of a specific wave, we would have to move in the same direction as the wave. When the wave looks stationary, our velocity is the same as that of the phase velocity of the wave.

If a wave is represented by $A \cos(kx - \omega t)$, where A is its amplitude, k the wave number, x the direction of propagation, ω the angular frequency and t the time, then following condition must be satisfied at this phase velocity because the wave looks stationary:

$$kx - \omega t = \text{constant} \quad \dots(1.36)$$

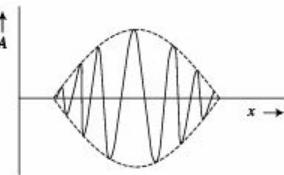


Fig. 1.14 A wave packet.

Differentiating, we get

$$\frac{dk}{dt} x + k \frac{dx}{dt} - \frac{d\omega}{dt} t - \omega = 0 \quad \dots(1.37)$$

If we assume that x and k are independent of time, the above equation becomes

$$k \frac{dx}{dt} - \omega = 0 \quad \dots(1.38)$$

or,

$$\frac{dx}{dt} = \frac{\omega}{k} = v_p \quad \dots(1.38)$$

where v_p is the phase velocity. Also,

$$v_p = \frac{\omega}{k} = \frac{2\pi\nu}{2\pi/\lambda} = v\lambda \quad \dots(1.39)$$

where λ is wavelength.

If the phase velocity is independent of frequency, a linear relation between ω and k is obtained as shown in Fig. 1.15 (a). However, an electron wave in a crystal does not meet this condition and the ω and k relationship becomes non-linear as shown in Fig. 1.15 (b). If we focus on a wave with a wave number of k_1 , we can define the velocity of this wave as

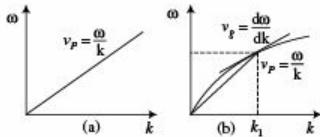


Fig. 1.15 (a) A linear ω versus k relationship and (b) A non-linear ω versus k relationship.

$$v_g = \frac{d\omega}{dk} \quad \dots(1.40)$$

where v_g is the group velocity. As $E = \hbar\omega$ also

$$v_g = \frac{1}{\hbar} \frac{dE(k)}{dk} \quad \dots(1.41)$$

Group velocity equals the particle velocity. For a particle having total energy E and potential energy V , we have kinetic energy

$$\frac{1}{2} mv^2 = E - V$$

or,

$$v = \sqrt{\frac{2(E-V)}{m}} \quad \dots(1.42)$$

According to de Broglie relation,

$$\lambda = \frac{h}{mv} = \frac{h}{m\sqrt{\frac{2(E-V)}{m}}}$$

or $\lambda = \frac{h}{m} \sqrt{\frac{m}{2(E-V)}}$... (1.43)

Group velocity $v_g = \frac{d\omega}{dk} = \frac{d(2\pi\nu)}{d(2\pi/\lambda)} = \frac{d\nu}{d\left(\frac{1}{\lambda}\right)}$

or, $\frac{1}{v_g} = \frac{d\left(\frac{1}{\lambda}\right)}{d\nu} = \frac{d}{d\nu} \left[\frac{m}{h} \sqrt{\frac{2(E-V)}{m}} \right]$
 $= \frac{1}{h} \frac{d}{d\nu} \sqrt{2m(h\nu - V)}$
 $= \frac{1}{h} \frac{1}{2} [2m(h\nu - V)]^{-1/2} (2mh)$

$\Rightarrow \frac{1}{v_g} = \frac{m}{\sqrt{2m(E-V)}} = \frac{1}{v}$... (1.44)

Thus, from Eqs. (1.42) and (1.44), we get

$$v_g = v$$

Hence, the motion of a material particle is equivalent to a group of waves or wave packet.

Relation Between Phase Velocity and Group Velocity

For a dispersive medium, there is variation of phase velocity with wavelength.

Phase velocity $v_p = \frac{\omega}{k}$

Group velocity $v_g = \frac{d\omega}{dk}$
 $= \frac{d\omega}{d\left(\frac{2\pi}{\lambda}\right)} = -\frac{\lambda^2}{2\pi} \frac{d\omega}{d\lambda}$
 $= -\frac{\lambda^2}{2\pi} \frac{d}{d\lambda} \left(\frac{2\pi v_p}{\lambda} \right) \quad \left(\because \omega = v_p k = \frac{2\pi}{\lambda} v_p \right)$
 $= -\lambda^2 \left(\frac{-v_p}{\lambda^2} + \frac{1}{\lambda} \frac{dv_p}{d\lambda} \right)$

or $v_g = v_p - \lambda \frac{dv_p}{d\lambda}$... (1.45)

Thus, the group velocity is smaller than the wave velocity. However, for a non-dispersive medium,

$$\frac{dv_p}{d\lambda} = 0 \text{ and thus } v_g = v_p.$$

1.14 UNCERTAINTY PRINCIPLE

According to Heisenberg's uncertainty principle, *it is impossible to measure both the exact position and exact momentum of an object at the same time.*

The limits of accuracy with which the position and momentum of particle can be obtained are specified by the uncertainty principle as

$$\Delta x \cdot \Delta p_x \geq h/4\pi \quad \dots(1.46)$$

or, $\Delta x \cdot \Delta p_x \geq \frac{\hbar}{2} \left(\text{as } \hbar = \frac{h}{2\pi} \right) \quad \dots(1.47)$

where, Δx is uncertainty in position of an object at some instant, Δp_x is uncertainty in the x component of the momentum at that instant. This equation states that the product of the uncertainty Δx in the position of an object at some instant and uncertainty Δp_x in its x component of the momentum at the same instant is equal to or greater than $h/4\pi$.

If we arrange matter so that Δx is small, corresponding to a narrow wave group, then Δp will be large. If we reduce Δp in some way, a broad wave group is inevitable and Δx will be large.

These uncertainties are not due to inaccurate apparatus but due to the inherent character in the nature of the quantities involved. Any instrumental or statistical uncertainty that arise during a measurement only increase the product $\Delta x \Delta p_x$. Since we cannot measure accurately both, where a particle is and what is its momentum, we cannot say anything definite about where it will be in future and how fast it will be moving then. We cannot know the future for sure because we cannot know the present for sure, but we can say that particle is more likely to be in one place than another and that its momentum is more likely to have a certain value than other.

Another equally useful form of uncertainty relation is in terms of energy and time,

$$\Delta E \cdot \Delta t \geq \frac{\hbar}{2} \quad \dots(1.48)$$

A more precise calculation based on the nature of wave groups gives

$$\Delta E \cdot \Delta t \geq \frac{\hbar}{2} \quad \dots(1.49)$$

It implies that the product of uncertainty ΔE in an energy measurement and the uncertainty Δt in the time at which the measurement is made is equal to or greater than $\frac{\hbar}{2}$.

In terms of angular momentum J and angle θ , the uncertainty principle is expressed as

$$\Delta J \cdot \Delta \theta \geq \frac{\hbar}{2} \quad \dots(1.50)$$

To gain further insight into the uncertainty relation and to understand the uncertainty principle better, a famous *thought experiment* single slit diffraction experiment is discussed below.

Single Slit Diffraction Experiment

Consider a beam of electrons travelling in the direction shown in Fig. 1.16. The slit AB of width Δx is perpendicular to the path of the electrons. Before entering the slit, the electron has a definite momentum $p = mv$. After passing through the slit the electron gets diffracted and acquires a momentum $p \sin \theta$. The angular deflection θ depends upon the component of the momentum parallel to the slit, i.e., $\Delta p = p \sin \theta = p\theta$, for small angular deflection.

The angle θ_0 in figure corresponds to the direction of first minimum of the diffraction pattern is $\theta_0 = \lambda/\Delta x$ for small value of θ_0 .

$$\Rightarrow \Delta x = \frac{\lambda}{\theta_0}$$

$$\therefore \Delta p \cdot \Delta x = p\theta \frac{\lambda}{\theta_0}$$

$$= p\lambda \times \frac{\theta}{\theta_0}$$

But $p\lambda = h$ and taking $\theta \approx \theta_0$

$$\Delta p \cdot \Delta x \sim h \left(\frac{\theta_0}{\theta_0} \right) \sim h$$

The probable deflection θ of the electron is less than θ_0 and according to Heisenberg, the uncertainty relation is given by:

$$\Delta p \cdot \Delta x \geq \frac{\hbar}{2}$$

Applications

(1) Non-existence of electrons and existence of protons and neutrons in nucleus: The radius of the nucleus of any atom is of the order of 10^{-14} m. If an electron is confined inside the nucleus, then uncertainty in the position Δx of the electron is equal to the diameter of the nucleus. That is, $\Delta x \approx 2 \times 10^{-14}$ m.

Using the uncertainty relation, the uncertainty in the momentum of the electron is given by

$$\Delta p_x \geq \frac{h}{2 \pi \Delta x} \geq \frac{1.055 \times 10^{-34}}{2 \times 10^{-14}}$$

$$\Delta p_x \geq 0.527 \times 10^{-20} \text{ Ns}$$

It means that the momentum component p_x and hence the magnitude of the total momentum $|\vec{p}|$ of the electron in the nucleus must be at least of the order of magnitude, i.e.,

$$|\vec{p}| \sim p_x \sim \Delta p_x \sim 0.527 \times 10^{-20} \text{ Ns}$$

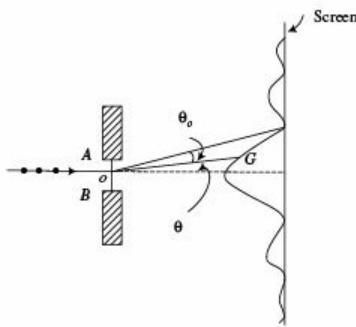


Fig. 1.16

Now the mass of the electron is 9.1×10^{-31} kg, the order of magnitude of the momentum (0.527×10^{-20} kg m/s) is relativistic and so

$$E^2 = p^2 c^2 + m_o^2 c^4$$

Neglecting $m_o^2 c^4$ (≈ 0.511 MeV) as it is much smaller than first term.

$$E = pc = (0.527 \times 10^{-20}) \times (3 \times 10^8) \text{ Joule}$$

$$= \frac{0.527 \times 3 \times 10^{-12}}{1.6 \times 10^{-19}} \text{ eV}$$

$$\approx 10 \text{ MeV}$$

Thus, if the electron exists inside the nucleus their energy must be of the order of 10 MeV. The electrons emitted during $\beta -$ decay in radioactive nuclei have energy of 3 to 4 MeV. Hence, in general, an electron cannot exist inside the nucleus.

For protons and neutrons, mass $m_o \approx 1.67 \times 10^{-27}$ kg

$$|v| = \frac{|\vec{p}|}{m_0} \sim 3 \times 10^6 \text{ m/s}$$

$$\therefore K.E. = \frac{p^2}{2m_0} = \frac{(0.527 \times 10^{-20})^2}{2 \times 1.67 \times 10^{-27}} \text{ J} \approx 52 \text{ keV}$$

Since the energy is less than the energies carried by these particles emitted by nucleus, protons and neutrons can exist inside the nucleus.

Binding Energy of an Electron in an Atom

In an atom, an electron is under the influence of electrostatic potential of the positively charged nucleus. It is confined within a linear dimension equal to the diameter of the electronic orbit. The uncertainty Δx is of the order of $2R$, R being the radius of the orbit and corresponding uncertainty in the momentum component Δp_x is given by

$$\Delta p_x \geq \frac{\hbar}{4\pi 2 R}$$

It shows that the momentum of an electron in an atomic orbit is at least

$$|\vec{p}| \sim p_x \sim \Delta p_x \sim \frac{\hbar}{4\pi 2 R} \sim 0.527 \times 10^{-24} \text{ kg ms}^{-1}$$

This is non-relativistic momentum for an electron.

$$\therefore K.E. = \frac{p^2}{2m_o} = \left(\frac{\hbar}{4\pi 2 R} \right)^2 \frac{1}{2m_o} = \frac{\hbar^2}{64\pi^2 m_o R^2}$$

$$P.E. \text{ of an electron} \quad V = -\frac{Ze^2}{4\pi\epsilon_o R}$$

$$\therefore E = K.E. + V = \frac{h^2}{64\pi^2 m_o R^2} - \frac{Ze^2}{4\pi\epsilon_0 R} \\ = \left[\frac{10^{-20}}{R^2} - \frac{15 \times 10^{-10} Z}{R^2} \right] \text{eV}$$

Taking $R = 10^{-10} \text{ m}$

$$E = (1 - 15 Z) \text{eV}$$

Binding energy of outermost electron in H and He are -13.6 eV and -24.6 eV respectively so the value of binding energy derived on the basis of uncertainty principle is acceptable as these are comparable in magnitudes.

(iv) Finite width of spectral lines: We know $\Delta E \cdot \Delta t \approx h$

Since life time of electron in an excited state is finite ($\approx 10^{-8} \text{ sec}$), so the energy levels of an atom

$$\Delta E = \frac{h}{\Delta t} \text{ must have a finite width}$$

It implies that excited levels must have a finite energy spread. Radiation given out by electron jump cannot be truly monochromatic. Thus, the spectral lines must have a natural finite width.

(v) Strength of Nuclear force: We know, nuclear radius, $r_o \approx 1.2 \times 10^{-13} \text{ cm}$

From uncertainty principle

$$p \sim \frac{\hbar}{r_o} \\ K.E. = \frac{p^2}{2m} \approx \frac{\hbar^2}{2mr_o^2} \approx 10 \text{ MeV} \quad (M - \text{Mass of proton})$$

Since the nucleus is bound so binding energy should be greater than kinetic energy with negative sign. So the binding energy of a nucleus is of the order of 10 MeV .

Example 1.16

Calculate the minimum uncertainty in the momentum of a ${}^4\text{He}$ atom confined to 0.40 nm .

Solution.

We know that ${}^4\text{He}$ atom is some where in the 0.40 nm region, therefore, $\Delta x = 0.40 \text{ nm}$. According to uncertainty relation

$$\Delta p_x \geq \frac{\hbar}{\Delta x}$$

For minimum uncertainty

$$\Delta p_x = \frac{\hbar}{\Delta x} = \frac{6.626 \times 10^{-34} \text{ Js}}{2\pi \times 0.40 \times 10^{-9} \text{ m}} \\ = 2.64 \times 10^{-25} \text{ kg m/s}$$

Example 1.17

An electron moves in the x -direction with a speed of 4×10^6 m/s. We can measure its speed to a precision of 1%. With what precision can its position simultaneously be measured?

Solution.

The momentum of the electron

$$\begin{aligned} p = mv &= (9.1 \times 10^{-31} \text{ kg}) (4 \times 10^6 \text{ ms}^{-1}) \\ &= 3.64 \times 10^{-24} \text{ kg ms}^{-1} \end{aligned}$$

$$\Delta p = \frac{3.64 \times 10^{-24}}{100} \text{ kg ms}^{-1} = 3.64 \times 10^{-26} \text{ kg ms}^{-1}$$

$$\begin{aligned} \Delta x &= \frac{\hbar}{2 \Delta p} = \frac{1.05 \times 10^{-34} \text{ Js}}{2 \times 3.64 \times 10^{-26}} \text{ kg ms}^{-1} \\ &= 1.44 \times 10^{-9} \text{ m} = 1.44 \text{ nm} \end{aligned}$$

Example 1.18

The average lifetime of an excited atomic state is 10^{-8} s. If the wavelength of the spectral line associated with a transition from this state is 4000 \AA , what is the width of this line?

Solution.

Average lifetime is 10^{-8} s

$$\Rightarrow \Delta t = 10^{-8} \text{ sec.}$$

$$E = h\nu = \frac{hc}{\lambda} \text{ or } |\Delta E| = \frac{hc}{\lambda^2} \Delta\lambda$$

$$(\Delta E) \cdot (\Delta t) \approx \frac{\hbar}{2} \text{ or } \frac{hc}{\lambda^2} (\Delta\lambda) (\Delta t) \approx \hbar/2$$

$$\begin{aligned} \Delta\lambda &= \frac{\lambda^2}{4\pi c(\Delta t)} = \frac{(4000 \times 10^{-10} \text{ m})^2}{4\pi(3 \times 10^8 \text{ ms}^{-1})10^{-8} \text{ s}} \\ &= 4.24 \times 10^{-15} \text{ m.} \end{aligned}$$

Example 1.19

The average lifetime of an excited atom is about 10^{-8} second. What is the order of natural width ($\Delta\nu$) of the line emitted by the atom?

Solution.

The order of natural line width is

$$\Delta\nu = \frac{1}{2\pi\Delta t} = \frac{10^8}{2\pi} \text{ Hz} = 1.6 \times 10^7 \text{ Hz.}$$

Example 1.20

The radius of an atomic nucleus is typically 5×10^{-15} m. What is the lowest limit of the energy that an electron must have to be in the atomic nucleus?

Solution.

The uncertainty in the electron's position is $\Delta x = 5 \times 10^{-15}$ m

$$\therefore \Delta p \geq \frac{\hbar}{\Delta x} = \frac{6.626 \times 10^{-34} \text{ Js}}{2\pi \times 5 \times 10^{-15} \text{ m}} = 2.11 \times 10^{-20} \text{ kg m/sec}$$

The momentum would also be of the same order if this is the uncertainty in it. This suggests that the kinetic energy of the electron is far greater than its rest energy and we can write

$$\begin{aligned} k.e. &= pc \quad \text{so that} \\ k.e. &= pc \geq (2.11 \times 10^{-20} \text{ kg ms}^{-1}) \times (3 \times 10^8 \text{ ms}^{-1}) \\ &\geq 6.33 \times 10^{-12} \text{ J} \\ &\geq 39 \text{ MeV} \end{aligned}$$

Thus, the *k.e.* of an electron must exceed 39 MeV for it to be a nuclear constituent. Experiments indicate that the electrons in an atom have only a fraction of this energy. Thus, we can conclude that electrons are not present in the atomic nuclei.

Example 1.21

The position x of a 0.01 g pellet has been carefully measured and is known within $\pm 0.5 \mu\text{m}$. What is the minimum uncertainty in its momentum?

Solution.

If x is known within $\pm 0.5 \text{ cm}$, the spread $\pm \Delta x$ in the position is certainly no longer than $0.5 \mu\text{m}$:

$$\Delta x \leq 0.5 \mu\text{m}$$

According to the uncertainty relation, this implies that the momentum is uncertain by an amount

$$\Delta p \geq \frac{\hbar}{2\Delta x} \geq \frac{10^{-34} \text{ Js}}{10^{-6} \text{ m}} = 10^{-28} \text{ kg m/s}$$

Example 1.22

Many excited states of atoms are unstable and decay by emission of a photon in a time of order $\Delta t = 10^{-8} \text{ s}$. What is the minimum uncertainty in the energy of such an atomic state?

Solution.

$$\begin{aligned} \text{According to the uncertainty relation, minimum uncertainty in energy is } \Delta E &\approx \frac{\hbar}{2\Delta t} = \frac{\hbar c}{2c\Delta t} \\ &= \frac{200 \text{ eV} \cdot \text{nm}}{2 \times (3 \times 10^{17} \text{ nm/s}) \times (10^{-8} \text{ s})} \\ &= 3 \times 10^{-8} \text{ eV} \end{aligned}$$

Example 1.23

An electron is known to be somewhere in an interval of total width $a \approx 0.1 \text{ nm}$ (the size of a small atom). What is minimum uncertainty in its velocity?

Solution.

If electron is certainly inside an interval of $\Delta x \leq a/2$

Since Δx is the spread from central value to either side.

According to uncertainty principle

$$\begin{aligned} \Delta p &\geq \frac{\hbar}{2\Delta x} \geq \frac{\hbar}{a} \\ \Rightarrow \Delta v &= \frac{\Delta p}{m} \geq \frac{\hbar}{am} \\ \Delta v &\geq \frac{\hbar c^2}{amc^2} = \frac{200 \text{ eV} \cdot \text{nm}}{(0.1 \text{ nm}) \times (0.5 \times 10^6 \text{ eV})} c \\ &= \frac{c}{250} = 10^6 \text{ m/s} \end{aligned}$$

Example 1.24

The uncertainty in the location of a particle moving with a velocity of $7.28 \times 10^7 \text{ m/s}$ is two times its de-Broglie wavelength. Find out the uncertainty in measuring the velocity.

Solution.

$$\text{de-Broglie wavelength} \quad \lambda = \frac{h}{mv}$$

$$\text{Given:} \quad \Delta x = 2\lambda = \frac{2h}{mv}$$

$$\text{According to uncertainty principle} \quad \Delta x \Delta p \approx \hbar$$

$$\Delta x \cdot m\Delta v \approx \hbar \quad \Rightarrow \quad \frac{2h}{mv} \cdot m\Delta v \approx \frac{h}{2\pi}$$

$$\text{or } \Delta v \approx \frac{v}{4\pi} = \frac{7.28 \times 10^7}{4 \times 3.14} = 5.79 \times 10^6 \text{ m/s}$$

1.15 WAVE FUNCTION

In water, wave that varies periodically is the height of the water surface. In sound waves, it is pressure. In light waves, electric and magnetic fields vary. What in the case of matter waves? It is *wave function* – ψ (psi). The value of the wave function associated with a moving body at the particular point x, y, z in space at the time t is related to the likelihood of finding the body at the time.

The wave function ψ has no direct physical significance. For the simple reason that ψ cannot be interpreted by an experiment. The probability that something be at a certain place at a given time must lie between 0 and 1. But the amplitude of any wave is negative as often as it is positive, and a negative probability, say –0.2, is meaningless. Hence, ψ by itself cannot be an observable quantity.

This objection does not apply to $|\psi|^2$, the square of the absolute value of the function, which is known as *probability density*.

The probability of experimentally finding the body described by the wave function ψ at the point x, y, z at the time t is proportional to the value of $|\psi|^2$ there at t .

A large value of $|\psi|^2$ means the strong possibility of the body's presence, while a small value of $|\psi|^2$ means the slight possibility of its presence. This interpretation was formulated by Born in 1926.

The quantity with which quantum mechanics is concerned is the wave function ψ of a body. While ψ itself has no physical interpretation, the square of its absolute amplitude $|\psi|^2$ evaluated at a particular place at a particular time is proportional to the probability of finding the body there at that time. The linear momentum, angular momentum and energy of the body are other quantities that can be obtained from ψ .

Wave functions are usually complex with both real and imaginary parts. A probability, however, must be a positive real quantity. The probability density $|\psi|^2$ for a complex ψ is therefore taken as the product $\psi^* \psi$ of ψ and its complex conjugate ψ^* . The complex conjugate of any function is obtained by replacing $i (= \sqrt{-1})$ by $-i$, wherever it appears in the function.

Wave function $\psi = A + iB$, where A and B are real function.

The complex conjugate ψ^* of ψ is $= A - iB$ and so $\psi^* \psi = A^2 - i^2 B^2 = A^2 + B^2$

Thus, $\psi^* \psi$ is always a positive real quantity.

Since $|\psi|^2$ is proportional to the probability density P of finding the body described by ψ , the integral of $|\psi|^2$ over all space must be finite—the body is somewhere, after all. If $\int_{-\infty}^{\infty} |\psi|^2 dv = 0$, the particle does not exist, and integral cannot be ∞ . Furthermore, $|\psi|^2$ cannot be negative or complex because of the way it is defined. Integral must be finite quantity if ψ is to describe properly a real body.

It is convenient to have $|\psi|^2$ equal to the probability density P of finding the particle described by ψ , rather than merely be proportional to P . If $|\psi|^2$ is equal to P , then it must be true that

$$\int_{-\infty}^{\infty} |\psi|^2 dz = 1 \text{ (Normalization)} \quad \dots(1.51)$$

Since if the particle exists somewhere at all times

$$\int_{-\infty}^{\infty} P dv = 1 \quad \dots(1.52)$$

A wave function that obeys Eq. (1.51) is said to be *normalized*. Every acceptable wave function can be normalized by multiplying it by an appropriate constant.

Besides being normalizable, ψ must be single valued, since P can have only one value at a particular place and time, and continuous. Momentum considerations require that the partial derivatives $\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial z}$, be finite, continuous, and single-valued. Thus, for a wave function to be correct wave function, it must satisfy the following conditions.

1. ψ must be continuous and single-valued everywhere.
2. $\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial z}$ must be continuous and single-valued everywhere.
3. ψ must be *normalizable*, which means that ψ must go to 0 as $x \rightarrow \pm\infty, y \rightarrow \pm\infty, z \rightarrow \pm\infty$ in order that $\int |\psi|^2 dv$ over all space be a finite constant.

The above rules are not always obeyed by the wave functions of particles in model situations that only approximate actual ones. For instance, the wave functions of a particle in a box with infinitely hard walls do not have continuous derivatives at the walls, since $\psi = 0$ outside the box. But in the real world, where walls are never infinitely hard, there is no sharp change in ψ at the walls and the derivatives are continuous.

Given a normalized and otherwise acceptable wave function ψ , the probability that the particle it describes will be in a certain region is simply the integral of the probability density $|\psi|^2$ over that region. Thus, for a particle restricted to motion in the x -direction, the probability of finding it between x_1 and x_2 is given by

$$P_{x_1 x_2} = \int_{x_1}^{x_2} |\psi|^2 dx \quad \dots(1.53)$$

1.16 BRA-KET NOTATION IN QUANTUM MECHANICS

Initially Werner Heisenberg, Max Planck and others formulated matrix representation in quantum mechanics, sometimes also called matrix mechanics. In matrix mechanics, a quantum state is represented by a quantum state vector and the physical quantities, such as position and momentum of a particle, are represented by matrices.

Later Paul Dirac developed bra-ket notation, which is concise and convenient of describing quantum states in matrix mechanics. He introduced and defined the symbol $|i\rangle$ to represent a quantum state i . This is called a *ket* or a *ket vector*.

Dirac notation uses a ket vector $|i\rangle$ to describe a quantum state instead of a wave function ψ_i , which is the basis of Schrödinger's notation. The complex conjugate of the wave function ψ_i^* is

represented by a bra vector $\langle i |$. The inner product of two complex functions can be described by the bra-ket notation as

$$\int \psi_i^* \psi_j dv = \langle \psi_i | \psi_j \rangle = \langle i | j \rangle \quad \dots(1.54)$$

If a wave function is normalized, the following must hold

$$\int \psi \psi^* dv = \int |\psi|^2 dv = \langle \psi | \psi \rangle = 1 \quad \dots(1.55)$$

A physical quantity or an action to determine physical quantities (measurements) is represented by A instead of a linear operator

$$A \psi = A|\psi\rangle \quad \dots(1.56)$$

The Schrödinger equation for a stationary (time independent) state is represented by

$$H_g |\psi_i\rangle = E_i |\psi_i\rangle \quad \dots(1.57)$$

where H_g is a Hamiltonian matrix. In matrix mechanics, a set of eigenvalues E_i can be obtained through the diagonalization of the Hamiltonian matrix instead of solving the Schrödinger equation. The principle of superposition is expressed as: if the two vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ satisfy the Schrödinger equation above, then a new vector $c_1|\psi_1\rangle + c_2|\psi_2\rangle$ also satisfies the same Schrödinger equation. In this state, the probability of observing an energy to be E_1 is $c_1 c_1^*$ and probability of observing the energy to be E_2 is $c_2 c_2^*$.

1.17 EXPECTATION VALUES

Once Schrödinger equation has been solved for a particle in a given physical situation, the resulting wave function $\psi(x, y, z, t)$ contains all information about the particle that is permitted by the uncertainty principle. Except for these variables that are quantized this information is in the form of probabilities and not specific numbers.

As an example, let us calculate the expectation value $\langle x \rangle$ of the position of a particle confined to the x -axis that is described by the wave function $\psi(x, t)$. This is the value of x we would obtain if we measured the positions of large number of particles described by the same wave function at some instant t and then averaged the results.

What is the average position \bar{x} of a number of identical particles distributed along the x axis in such a way that there are N_1 particles at x_1 , N_2 particles at x_2 , and so on? The average position in this case is the same as the centre of mass of the distribution, and so

$$\bar{x} = \frac{N_1 x_1 + N_2 x_2 + N_3 x_3 + \dots}{N_1 + N_2 + N_3} = \frac{\sum x_i N_i}{\sum N_i}$$

When we are dealing with a single particle, we must replace the number N_i of particles at x_i by the probability P_i that the particle be found in an interval dx at x_i . This probability is

$$P_i = |\psi_i|^2 dx \quad \dots(1.58)$$

where ψ_i is the particle wave function evaluated at $x = x_i$. Making this substitution and changing the summations to integrals, we see that the expectation value of the position of the single particle is

$$\langle x \rangle = \frac{\int_{-\infty}^{\infty} x |\psi|^2 dx}{\int_{-\infty}^{\infty} |\psi|^2 dx} \quad \dots(1.59)$$

If ψ is a normalized wave function, the denominator of Eq. (1.59) equals the probability that the particle exists somewhere between $x = -\infty$ and $x = \infty$ and therefore has the value 1. In this case

$$\text{Expectation value for position } \langle x \rangle = \int_{-\infty}^{\infty} x |\psi|^2 dx \quad \dots(1.60)$$

This formula states that $\langle x \rangle$ is located at the centre of mass (so to speak) of $|\psi|^2$. If $|\psi|^2$ is plotted versus x on a graph and the area enclosed by the curve and the x -axis is cut out, the balance point will be at $\langle x \rangle$.

The same procedure as that followed above can be used to obtain the expectation value $\langle G(x) \rangle$ of any quantity for instance, potential energy $U(x)$ – that is, a function of the position x of a particle described by a wave function ψ . The result is

$$\text{Expectation value } \langle G(x) \rangle = \int_{-\infty}^{\infty} G(x) |\psi|^2 dx \quad \dots(1.61)$$

The expectation value $\langle p \rangle$ of momentum cannot be calculated this way because, according to the uncertainty principles, no such function as $p(x)$ can exist. If we specify x so that $\Delta x = 0$, we cannot specify a corresponding p since $\Delta x \Delta p \geq \frac{\hbar}{2}$.

The same problem occurs for the expectation value $\langle E \rangle$ for energy.

1.18 OPERATORS IN QUANTUM MECHANICS

In quantum mechanics, physical quantities or actions undertaken to determine physical quantities (measurements) are represented by linear operators. We may define *an operator as a rule by which, given any function, we can find another function*. Thus, we can define an operator P as: multiply the function by an independent variable, i.e., $Pf(x) = xf(x)$. Similarly, we can define an operator Q as the differentiation w.r.t. the independent variable, i.e., $Qf(x) = f'(x)$. The result obtained by the action of P followed by Q is different from that obtained by the action of Q followed by P as shown below.

$$PQ f(x) = P[Qf(x)] = xf'(x) \quad \dots(1.62)$$

$$QP f(x) = Q[Pf(x)] = xf'(x) + f(x) \quad \dots(1.63)$$

For the two operators α and β , if $\alpha\beta = \beta\alpha$, α and β are said to **commute** or be **commutable**. In this sense, the operators P and Q are not commutable.

A measurement changes the state of a particle represented by a wavefunction (i.e., an operator operating on the wave function). Suppose, we measure the physical quantities a and b of an electron by continuous measurements A and B . The result obtained by carrying out first measurement A

followed by the measurement B is different from that obtained by conducting first measurement followed by measurement A , because each of the measurements changes the wave function. This can be represented mathematically using operators acting on the wave function as

$$AB \psi \neq BA \psi \quad \dots(1.64)$$

This equation indicates that the operators A and B are not commutable. In this case, measurements change the wave function and the two physical quantities corresponding to the operators A and B cannot be determined precisely at the same time due to the laws defined by uncertainty principle. On the other hand, if the operators A and B are commutable, we can determine the two physical constants at the same time without the restrictions of the uncertainty principle.

Following are the operators corresponding to some important physical quantities:

<i>Physical quantities</i>	<i>Symbol</i>	<i>Operator</i>
Position	$q(\text{or } r)$	$q(\text{or } x, y, z; r)$
Energy	E	$i\hbar \frac{\partial}{\partial t}$
Kinetic energy	T	$-\hbar^2 \frac{\nabla^2}{2m}$
Momentum	p	$-i\hbar \nabla$ $(-i\hbar \frac{\partial}{\partial x}, -i\hbar \frac{\partial}{\partial y}, -i\hbar \frac{\partial}{\partial z}, \text{ for } x, y$ and z -components respectively)
Velocity	v	$-\frac{i\hbar}{m} \nabla$ $\left(\frac{-i\hbar}{m} \frac{\partial}{\partial x}, \frac{-i\hbar}{m} \frac{\partial}{\partial y}, \frac{-i\hbar}{m} \frac{\partial}{\partial z} \right)$
Hamiltonian	H	$i\hbar \frac{\partial}{\partial t}$

For example, Schrödinger equation

$$H\psi = E\psi$$

can be expressed as

$$H\psi = i\hbar \frac{\partial \psi}{\partial t} \text{ or } (H)\psi = \left(i\hbar \frac{\partial}{\partial t} \right) \psi$$

where $\hbar = \frac{h}{2\pi}$ and h is Planck's constant.

Similarly, if P and Q are the position and momentum operators respectively, the $PQ\psi$ can be represented as

$$PQ \psi = -i\hbar \frac{\partial}{\partial q} (q \psi) = -i\hbar\psi - i\hbar q \frac{\partial\psi}{\partial q} \quad \dots(1.65)$$

Similarly, $QP\psi = q \left(-i\hbar \frac{\partial\psi}{\partial q} \right) = -i\hbar q \frac{\partial\psi}{\partial q}$... (1.66)

Thus, we see that

$$(PQ - QP)\psi = -i\hbar\psi \neq 0 \quad \dots(1.67)$$

1.19 SCHRÖDINGER EQUATION: TIME DEPENDENT FORM

In quantum mechanics the wave function ψ corresponds to the wave variable y of wave motion in general. However, ψ , unlike y , is not itself a measurable quantity and may therefore be complex. For this reason, we assume that ψ for a particle moving freely in the $+x$ direction is specified by

$$\psi = A e^{-i\omega(t-x/v)} \quad \dots(1.68)$$

Replacing ω by $2\pi\nu$ and v by λv , we get

$$\psi = A e^{-2\pi i \left(vt - \frac{x}{\lambda} \right)} \quad \dots(1.69)$$

This is convenient since we already know that what v and λ are in terms of the total energy E and momentum p of the particle being described by ψ because

$$E = h\nu = 2\pi\hbar\nu \text{ and } \lambda = \frac{h}{p} = \frac{2\pi\hbar}{p}$$

\therefore For free particle $\Rightarrow \psi = A e^{-(i/\hbar)(Et - px)}$... (1.70)

Equation (1.70) describes the wave equivalent of an unrestricted particle of total energy E and momentum p moving in the $+x$ direction.

The expression for the wave function ψ given by Eq. (1.70) is correct only for freely moving particles. However, we are most interested in situations where the motion of a particle is subjected to various restrictions. An important concern, for example, is an electron bound to an atom by the electric field of its nucleus. What we must now do is to obtain the fundamental differential equation for ψ , which we can then solve for ψ in a specific situation. This equation, called Schrödinger equation, can be arrived at in various ways, but it cannot be rigorously derived from existing physical principles. The equation represents something new. What will be done here is to show one route to the wave equation for ψ and then to discuss the significance of the result.

We begin by differentiating Eq. (1.70) for ψ twice w.r.t. x , giving

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{p^2}{\hbar^2} \psi \Rightarrow p^2\psi = -\hbar^2 \frac{\partial^2 \psi}{\partial x^2} \quad \dots(1.71)$$

Differentiating Eq. (1.70) w.r.t. t gives

$$\frac{\partial \psi}{\partial t} = -\frac{iE}{\hbar} \psi \Rightarrow E\psi = -\frac{\hbar}{i} \frac{\partial \psi}{\partial t} \quad \dots(1.72)$$

At speeds, small compared with that of light, the total energy E of a particle is the sum of its kinetic energy $\frac{p^2}{2m}$ and its potential energy U , where U is in general a function of position x and time t , i.e.,

$$E = \frac{p^2}{2m} + U(x, t) \quad \dots(1.73)$$

The function U represents the influence of the rest of the universe on the particle, of course, only a small part of the universe interacts with the particle to any extent. For instance, in the case of the electron in a hydrogen atom, only the electric field of the nucleus must be taken into account.

Multiplying both sides of Eq. (1.73) by ψ , we get

$$E\psi = \frac{p^2\psi}{2m} + U\psi \quad \dots(1.74)$$

Substituting for $E\psi$ and $p^2\psi$ from Eq. (1.72) and (1.71), we get *time dependent form of Schrödinger equation* in one dimension

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + U\psi \quad \dots(1.75)$$

In 3-D, the time-dependent form of Schrödinger equation is

$$i\hbar \frac{\partial \psi}{\partial t} = \frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) + U\psi \quad \dots(1.76)$$

where the particle's potential energy U is some function of x, y, z and t .

Any restrictions that may be present on the particle's motion will affect the potential energy function U . Once U is known, Schrödinger equation may be solved for the wave function ψ of the particle, from which its probability density $|\psi|^2$ may be determined for a specific x, y, z, t .

Schrödinger equation is a basic principle itself and cannot be derived from other basic principles of physics.

Schrödinger equation is remarkably accurate in predicting the results of experiments for atomic and subatomic particles.

1.20 SCHRÖDINGER'S EQUATION: STEADY STATE FORM (TIME INDEPENDENT)

In many situations, the potential energy of a particle does not depend on time explicitly, the forces that act on it, and hence U , vary with the position of the particle only. When this is true, Schrödinger equation may be simplified by removing all reference to t .

One-dimensional wave function ψ of an unrestricted particle may be expressed as

$$\Psi(t) = A e^{-(iE/\hbar)t} (Et - px) = A e^{-i(E/\hbar)t} e^{+(ip/\hbar)x} = \psi e^{-(iE/\hbar)t} \quad \dots(1.77)$$

where Ψ is the product of a time-dependent function $e^{-(iE/\hbar)t}$ and a position dependent function ψ . As it happens, the time variations of all wave functions of particles acted on by stationary forces have the same form as that of an unrestricted particle.

Substituting the Ψ of Eq. (1.77) into the time-dependent form of Schrödinger equation

$$\left(i\hbar \frac{\partial \psi}{\partial t} - \frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V\psi \right), \text{ we find that}$$

$$E\psi e^{-(iE/\hbar)t} = -\frac{\hbar^2}{2m} e^{-(iE/\hbar)t} \frac{\partial^2 \psi}{\partial x^2} + U\psi e^{-(iE/\hbar)t}$$

Dividing throughout by the common exponential factor gives

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{2m}{\hbar^2} (E - U) \psi = 0 \quad \dots(1.78)$$

Equation (1.78) is the *steady state form of Schrödinger's equation*. In 3-D the corresponding Schrödinger equation is

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{2m}{\hbar^2} (E - U) \psi = 0 \quad \dots(1.79)$$

$$\text{As, } \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \nabla^2 \Rightarrow \nabla^2 \psi + \frac{2m}{\hbar^2} (E - U) \psi = 0 \quad \dots(1.80)$$

An important property of Schrödinger's steady state equation is that, if it has one or more solutions for a given system, each of these wave functions corresponds to a specific value of the energy E . Thus, *energy quantization appears in wave (or quantum) mechanics as a natural element of the theory*, and energy quantization in the physical world is revealed as a universal phenomenon characteristic of all stable systems.

1.21 EIGENVALUES AND EIGENFUNCTIONS

The values of energy E_n for which Schrödinger's steady state equation can be solved are called *Eigenvalues* and the corresponding wave function ψ_n are called *Eigenfunctions*. (These terms come from the German Eigenwert meaning proper or characteristic value and Eigenfunction meaning proper or characteristic function). The discrete energy levels of the hydrogen atom

$$E_n = -\frac{me^4}{32\pi^2 \epsilon_0^2 \hbar^2} \left(\frac{1}{n^2} \right), \quad n = 1, 2, 3, \dots \quad \dots(1.81)$$

are an example of a set of eigenvalues.

An important example of a dynamical variable other than total energy that is found to be quantized in stable systems is angular momentum L . In the case of the hydrogen atom, the eigenvalues of the magnitude of the total angular momentum are specified by

$$L = \sqrt{l(l+1)\hbar} \quad l = 0, 1, 2, \dots, (n-1) \quad \dots(1.82)$$

1.22 DEGENERATE AND NON-DEGENERATE EIGENFUNCTIONS

When $H\psi_n = E_n\psi_n$ is solved, corresponding to each eigenvalue there is only one eigenfunction. In such a situation, the eigenfunction is known as **degenerate**.

Eigenfunction: However, there are situations when corresponding to a single eigenvalue, there are a number of eigenfunctions, then these eigenfunctions are known as *degenerate eigenfunctions*.

In quantum mechanics, there are situations when the eigenvalue depends on principal quantum number n (E_n), however, eigenfunction also depends upon orbital quantum number l and magnetic quantum number m (ylm). As for each n , l can vary from 0 to $n - 1$ and for each of these l values, m can vary from $-l$ to $+l(2l + 1)$ values). This leads to degeneracy. The total degeneracy is

$$\sum_{l=0}^{n-1} (2l+1) = \frac{2(n)(n-1)}{2} + n = n^2.$$

For $n = 2$, $l = 0$, $m = 0$; $l = 1$, $m = 0$; $l = 1$, $m = \pm 1$, giving four wave functions, i.e., system becomes fourfold degenerate. Similarly, you can work out for $n = 3$, system becomes nine fold degenerate and for $n = 4$, the system becomes sixteen fold degenerate, and so on.

1.23 PARTICLE IN A BOX

The simplest quantum-mechanical problem is that of a particle in a box with infinitely hard walls. We may specify the particle's motion by saying that it is restricted to travel along the x -axis between $x = 0$ and $x = L$ by infinitely hard walls (Fig. 1.17). A particle does not lose energy when it collides with such walls, so that its total energy stays constant. From a formal point of view the potential energy U of the particle is infinite on both sides of the box, while U is a constant-say zero for convenience, inside the box. Because the particle cannot have an infinite amount of energy, it cannot exist outside the box and so its wave function ψ is 0 for $x \leq 0$ and $x \geq L$. Our task is to find what ψ is within the box, namely, between $x = 0$ and $x = L$.

Within the box Schrödinger equation becomes

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} E \psi = 0 \quad \dots(1.83)$$

Since $U = 0$ inside the box. (The total derivative $\frac{d^2\psi}{dx^2}$ is the same as the partial derivative $\frac{\partial^2\psi}{\partial x^2}$

because ψ is a function only of x in this problem) Equation (1.83) has the general solution

$$\psi = A \sin \frac{\sqrt{2mE}}{\hbar} x + B \cos \frac{\sqrt{2mE}}{\hbar} x \quad \dots(1.84)$$

which we can verify by substitution back into (1.83), A and B are constants to be evaluated.

This solution is subject to the boundary conditions that $\psi = 0$ for $x = 0$ and for $x = L$. Since $\cos 0 = 1$, the second term cannot describe the particle because it does not vanish at $x = 0$. Hence, we conclude that $B = 0$. Since $\sin 0 = 0$, the sine term always yields $\psi = 0$ at $x = 0$, as required, but ψ will be zero at $x = L$ only when

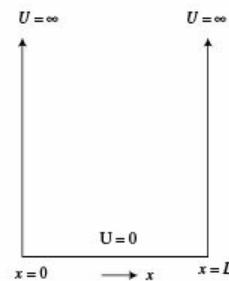


Fig. 1.17 A potential well with infinitely high barriers a box with infinitely hard walls

$$\frac{\sqrt{2mE}}{\hbar}L = n\pi, \quad n = 1, 2, 3, \dots \quad \dots(1.85)$$

This is because the sines of the angles $\pi, 2\pi, 3\pi, \dots$ are all 0. From Eq. (1.85) it is clear that the energy of the particle can have only certain values which are the eigenvalues. These eigenvalues, constituting the energy levels of the system, are found by solving Eq. (1.85) for E_n , which gives

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}, \quad n = 1, 2, 3, \dots \quad \dots(1.86)$$

Wave Functions of A Particle in a Box

The wave functions of a particle in a box whose energies are E_n , from Eq. (1.84) with $B = 0$, are

$$\psi_n = A \sin \frac{\sqrt{2mE_n}}{\hbar}x \quad \dots(1.87)$$

Substituting Eq. (1.86) for E_n gives

$$\psi_n = A \sin \sqrt{\frac{2m}{2mL^2} \frac{n^2 \pi^2 \hbar^2}{\hbar}} x = A \sin \frac{n\pi x}{L} \quad \dots(1.88)$$

For the eigenfunctions corresponding to the energy eigenvalues E_n ,

These eigenfunctions meet all the requirements. For each quantum number n , ψ_n is a finite, single-valued function of x , and ψ_n and $\frac{\partial \psi_n}{\partial x}$ are continuous (except at the ends of the box).

Furthermore, the integral of $|\psi_n|^2$ over all space is finite, as we can see by integrating $|\psi_n|^2$ from $x = 0$ to $x = L$ (since the particle is confined within these limits). With the help of the trigonometric identity

$$\sin^2 \theta = \frac{1}{2} (1 - \cos 2\theta)$$

we find that

$$\begin{aligned} \int_{-\infty}^{\infty} |\psi_n|^2 dx &= \int_0^L |\psi_n|^2 dx = A^2 \int_0^L \sin^2 \left(\frac{n\pi x}{L} \right) dx = \frac{A^2}{2} \left[\int_0^L dx - \int_0^L \cos \left(\frac{2n\pi x}{L} \right) dx \right] \\ &= \frac{A^2}{2} \left[x - \left(\frac{L}{2n\pi} \right) \sin \frac{2n\pi x}{L} \right]_0^L = A^2 \left(\frac{L}{2} \right) \end{aligned} \quad \dots(1.89)$$

To normalize Ψ we must assign a value to A such that $|\psi_n|^2 dx$ is equal to the probability Pdx of finding the particle between x and $x + dx$, rather than merely proportional to Pdx . If $|\psi_n|^2 dx$ is equal Pdx , then it must be true that

$$\int_{-\infty}^{\infty} |\psi|^2 dx = 1 \quad \dots(1.90)$$

Comparing Eqs. (1.89) and (1.90), we see that the wave function of a particle in a box are normalized if

$$A = \sqrt{\frac{2}{L}} \quad \dots(1.91)$$

The normalized wave functions of the particle are therefore

$$\Psi_n = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}, \quad n = 1, 2, 3, \dots \quad \dots(1.92)$$

The normalized wave function ψ_1 , ψ_2 and ψ_3 together with the probability density $|\psi_1|^2$, $|\psi_2|^2$ and $|\psi_3|^2$ are plotted in Fig. 1.18. Although ψ_n may be negative as well as positive, $|\psi_n|^2$ is always +ve and since ψ_n is normalized, its value at a given x is equal to the probability density of finding particle there. In every case $|\psi_n|^2 = 0$ at $x = 0$ and $x = L$, the boundaries of the box.

At a particular place in the box the probability of the particle being present may be very different for different quantum numbers. For instance, $|\psi_1|^2$ has its maximum value of $\frac{2}{L}$ in the middle of the box, while a particle in the next higher state of $n = 2$ is never there.

The wave function shown in Fig. 1.18(a) resemble the possible vibrations of a string fixed at both ends. This follows from the fact that waves in a stretched string and the wave representing a moving particle are described by equations of the same form, so that when identical restrictions are placed upon each kind of wave, and therefore the formal results are identical.

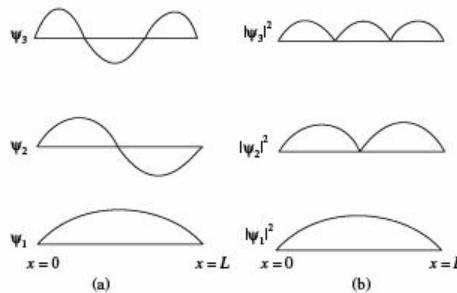


Fig. 1.18 (a) Wave functions and (b) Probability densities.

1.24 SIMPLE HARMONIC OSCILLATOR

Simple harmonic oscillator is of fundamental importance in physics because there are many quantum systems which can be approximated by harmonic oscillators. For example, the variation of diatomic molecules and motions of atoms in a crystal lattice can be treated to a first approximation as the motion of particles in harmonic fields. We know that the potential energy $U(x)$ of a harmonic oscillator of mass m is given by

$$U(x) = \frac{1}{2} kx^2 \quad \dots(1.93)$$

where $k (= m \omega^2)$ is the force (or spring) constant defined as the force per unit displacement. The Schrödinger equation for this system can be written as

$$\frac{d^2 \Psi}{dx^2} + \frac{2m}{\hbar^2} \left(E - \frac{k}{2} x^2 \right) \Psi = 0 \quad \dots(1.94)$$

The solution of the Schrödinger equation yields the following wave functions

$$\begin{aligned} \Psi_n &= A_n H_n(z) e^{-z^2/2} \\ &= A_n H_n \left[\left(\frac{mk}{\hbar^2} \right)^{\frac{1}{2}} x \right] \exp \left[-\frac{1}{2} \left(\frac{mk}{\hbar^2} \right)^{\frac{1}{2}} x^2 \right] \end{aligned} \quad \dots(1.95)$$

and eigenvalues

$$\begin{aligned} E_n &= \hbar \sqrt{\frac{k}{m}} \left(n + \frac{1}{2} \right) \\ &= \left(n + \frac{1}{2} \right) \hbar \omega, \quad n = 0, 1, 2, \dots \end{aligned} \quad \dots(1.96)$$

where A_n and $H_n(z)$ (Hermite polynomials) are given by

$$A_n = \left[\frac{(mk/\hbar)^{\frac{n}{2}}}{2^n n! \pi^{\frac{1}{2}}} \right]^{\frac{1}{2}}$$

$$H_n(z) = (-1)^n \exp(z^2) \frac{d^n}{dz^n} \exp(-z^2) \quad \dots(1.97)$$

The first four energy levels, corresponding wave functions $\Psi_n(x)$ and probability densities $|\Psi_n(x)|^2$ are shown in Fig. 1.19 (b) and (c) respectively. It is important to note from the figures that $\Psi_n(x)$ and $|\Psi_n(x)|^2$ have a finite value outside the classically allowed region. The probability density oscillates between the turning points and decreases exponentially beyond the turning points. The figure clearly indicates the penetration of the particle into the region outside the classical turning points.

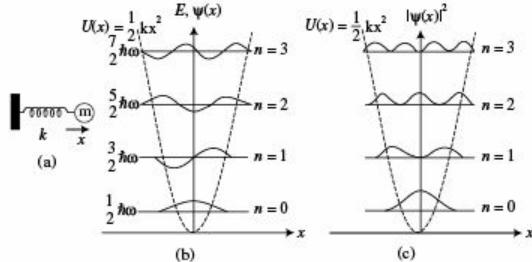


Fig. 1.19 (a) Harmonic oscillator, (b) energy levels and corresponding wave function, and (c) corresponding probability densities.

The energy levels (given by Eq. 1.96) are equally spaced, as shown in Fig. 1.19 (b). For the ground state ($n = 0$), from Eq.(1.96), we have

$$E_o = \frac{1}{2} \hbar \omega$$

Thus, the oscillator will have this energy even at absolute zero. This energy is referred to as **zero point energy**. This is in contradiction with Planck oscillator for which $E_n = n\hbar\omega$ and so for $n = 0$, $E_0 = 0$.

It is important to realize that the existence of zero point energy is in accordance with the uncertainty principle. The laws of quantum mechanics do not permit the harmonic oscillator to have zero energy and the smallest allowed energy is the zero point energy. If the oscillator has zero energy, it will also have zero momentum and therefore can be located exactly at the position of minimum potential energy, which will not be in accordance with the uncertainty principle. Hence, the necessary uncertainties in position and momentum give rise to the zero point energy.

A comparison of energy values ($E_n = \left(n + \frac{1}{2}\right)\hbar\omega$ vs $E_n = n\hbar\omega$) makes it evident that the

quantum mechanical values are higher by $\frac{1}{2}\hbar\omega$. Also, since $H_0(z) = 1$, the ground state wave function can be shown to be

$$\Psi_o = \left(\frac{m\omega}{\hbar\pi}\right)^{\frac{1}{4}} \exp\left(-\frac{m\omega x^2}{2\hbar}\right) \quad \dots(1.98)$$

Example 1.25

The strongest vibrational bond of CO molecule occurs at 6.43×10^{13} Hz. If the reduced mass of CO is 1.1385×10^{-26} kg, calculate (i) the force constant and (ii) the zero point energy of the CO bond.

Solution.

Due to the vibrational motion of the C and O atoms, the C – O bond stretches, then contracts and so on, i.e., undergo simple harmonic motion.

(i) We know $v_o = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$, where μ is the reduced mass.

$$\begin{aligned} \therefore \text{force constant} \quad k &= 4\pi^2 v_o^2 \mu \\ &= 4 \times (3.14)^2 (6.43 \times 10^{13})^2 (1.1385 \times 10^{-26}) \\ &= 1860 \text{ Nm}^{-1} \end{aligned}$$

$$\begin{aligned} \text{(ii) Zero point energy} \quad E_o &= \frac{1}{2} \hbar v_o \\ &= \frac{1}{2} \times 6.62 \times 10^{-34} \times 6.43 \times 10^{13} \text{ J} \\ &= 21.3 \times 10^{-21} \text{ J} \end{aligned}$$

$$= \frac{21.3 \times 10^{-21}}{1.6 \times 10^{-19}} \text{ eV}$$

$$= 0.133 \text{ eV.}$$

1.25 RECTANGULAR POTENTIAL BARRIER:TUNNEL EFFECT

Let us consider transmission of a plane wave through a rectangular barrier. The transmission is based on tunnel effect, according to which if a particle is impinging on a barrier with energy less than the height of the potential barrier, it will not necessarily be totally reflected by the barrier but there is always a probability that it may cross the barrier and continue its forward motion.

Consider a particle incident on the barrier of the type shown in Fig. 1.20, with the condition that the energy of particle $E < U_0$ and we have the different potentials in different regions as given below.

$$\begin{aligned} U(x) &= 0 && \text{for } x < 0 \text{ in region I} \\ U(x) &= U_0 && \text{for } 0 < x < a \text{ in region II} \\ U(x) &= 0 && \text{for } x > a \text{ in region III.} \end{aligned}$$

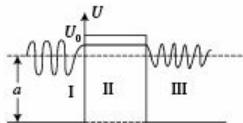


Fig. 1.20

According to classical mechanics, a particle striking an infinitely hard wall has no chance of getting through it. Quantum mechanics is in agreement as a particle of finite kinetic energy can not enter a region where the potential energy is $U = \infty$.

But what about a less formidable wall, one that is not infinitely hard but still requires more energy U to penetrate than the energy E ? In this case, classical mechanics requires the particle to bounce off but now quantum mechanics gives a different result, there is a certain likelihood not necessarily high, but not zero either, that the particle can pass through the energy barrier even though $E < U$. Although the particle does not have enough energy to surmount the barrier, it can nevertheless **tunnel** through it. The higher the barrier and the thicker it is, the smaller the chance that the particle can get through.

The tunnel effect can be understood in terms of the uncertainty principle; if we say that the incident particle cannot enter the barrier, then the uncertainty Δx in its position must be zero there.

But since $\Delta x \Delta p \geq \frac{\hbar}{2}$, the corresponding uncertainty Δp in the

particle's momentum must be infinite inside the barrier. An infinite uncertainty in p means that p and E as well, must be infinite, which is not compatible with the finite momentum and energy of the particle. Therefore, the particle must be able to enter the barrier, and once inside it, has the possibility of continuing on.

The tunnel effect actually occurs, notably in the case of α -particles emitted by certain radioactive nuclei. An α -particle whose kinetic energy is only a few MeV is able to escape from a nucleus

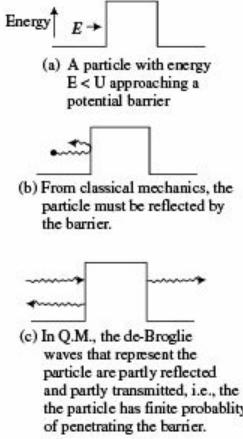


Fig. 1.21

whose potential wall is as high as 25 MeV. The probability of escape is so small that the α -particle might have to strike the wall 10^{38} or more times before it emerges but sooner or later it does get out. Tunnelling also occurs in the operation of certain semiconductor diodes in which electrons pass through the potential barriers even though their kinetic energies are smaller than the barrier heights.

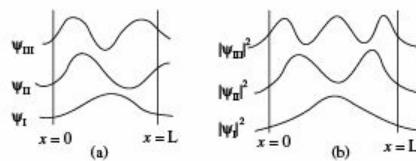


Fig. 1.22 (a) Wave functions (b) Probability densities of a particle in a finite potential well. The particle has a certain probability of being outside the well.

Figure 1.22 shows the first three wave functions and the corresponding probability densities for a particle in a finite potential well. The wave functions do not equal zero outside the well, as they do for a particle in an infinite potential well. Even though the particle's energy is less than the value of U outside the well, there is still a definite chance that it will be found there. The particle wavelengths that can fit into the well are accordingly somewhat longer than that in the case of the infinite potential well, which means lower momenta and hence lower energy levels.

Tunnel Effect

Let us consider a beam of identical particles which have the kinetic energy $K = E$ ($E < U$). The beam is incident from the left on a potential barrier of height U and width L , as shown in Fig. 1.23. On both sides of the barrier $U = 0$, which means that no forces act on the particles there. In these regions Schrödinger equation for the particles takes the form

$$\frac{d^2 \Psi_I}{dx^2} + \frac{2m}{\hbar^2} E \Psi_I = 0 \quad \dots(1.99)$$

$$\frac{d^2 \Psi_{\text{III}}}{dx^2} + \frac{2m}{\hbar^2} E \Psi_{\text{III}} = 0 \quad \dots(1.100)$$

The general solutions of these equations are

$$\Psi_I = A e^{ik_1 x} + B e^{-ik_1 x} \quad \dots(1.101)$$

$$\Psi_{\text{III}} = F e^{ik_1 x} + G e^{-ik_1 x} \quad \dots(1.102)$$

where, wave number outside barrier $k_1 = \sqrt{\frac{2mE}{\hbar^2}} = \frac{P}{\hbar} = \frac{2\pi}{\lambda}$ is the wave number of the de-Broglie

waves that represent the particles outside the barrier. We know

$$e^{i\theta} = \cos \theta + i \sin \theta \text{ and } e^{-i\theta} = \cos \theta - i \sin \theta$$

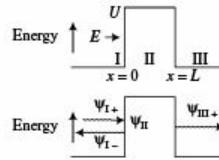


Fig. 1.23

$$\dots(1.103)$$

These solutions are equivalent to equation $\psi = A \sin \frac{\sqrt{2mE}}{\hbar}x + B \cos \frac{\sqrt{2mE}}{\hbar}x$. The values of the coefficients are different in each case.

Incident and Reflected Waves

The various terms in Eqs. (1.101) and (1.102) are not hard to interpret. As shown in Fig. 1.23, Ae^{ik_1x} is a wave of amplitude A incident from the left on the barrier. Hence, we can write the incoming wave as

$$\Psi_{I+} = Ae^{ik_1x} \quad \dots(1.103)$$

which corresponds to the incident beam of particles in the sense that $|\Psi_{I+}|^2$ is their probability density. If v_{I+} is the group velocity of the incoming wave, which equals the velocity of the particles, then

$$S = |\Psi_{I+}|^2 v_{I+} \quad \dots(1.104)$$

is the flux of particles that arrive at the barrier. That is, S is the *number of particles per unit area per second* that arrive there. At $x = 0$, the incident wave strikes the barrier and is partially reflected. The reflected wave is given by

$$\Psi_{I-} = Be^{-ik_1x} \quad \dots(1.105)$$

Hence,

$$\Psi_I = \Psi_{I+} + \Psi_{I-} \quad \dots(1.106)$$

On the other side of the barrier ($x > L$) there can only be one wave i.e., transmitted wave

$$\Psi_{III} = Fe^{ik_3x} \quad \dots(1.107)$$

travelling in the $+x$ direction at the velocity v_{III+} , since region III contains nothing that could reflect the wave. Hence, $G = 0$ and

$$\Psi_{III} = \Psi_{III+} = Fe^{ik_3x} \quad \dots(1.108)$$

The *transmission probability* T for a particle to pass through the barrier is the ratio

$$T = \frac{|\Psi_{III+}|^2 v_{III+}}{|\Psi_{I+}|^2 v_{I+}} = \frac{FF^* v_{III+}}{AA^* v_{I+}} \quad \dots(1.109)$$

between the flux of particles that emerges from the barrier and the flux that arrives at it. In other words, T is the fraction of incident particles that succeed in tunnelling through the barrier. Classically $T = 0$ because a particle with $E < U$ cannot exist inside the barrier. Let us investigate the Quantum mechanical possibility.

In region II, Schrödinger equation for the particles is

$$\frac{d^2 \Psi_{II}}{dx^2} + \frac{2m}{\hbar^2}(E - U)\Psi_{II} = 0 \quad \dots(1.110)$$

Since $U > E$, its general solution

$$\Psi_{II} = C e^{ik_2x} + D e^{-ik_2x} \quad \dots(1.111)$$

where $k' = \frac{\sqrt{2m(E-U)}}{\hbar}$

Since $E < U$, k' is imaginary and it is convenient to define another wave number k_2 in the following ways: wave number inside the barriers

$$k_2 = -ik' \quad \dots(1.112)$$

In terms of k_2 , Eq. (1.111) is given by

$$\psi_{II} = Ce^{-k_2 x} + De^{k_2 x} \quad \dots(1.113)$$

Since the exponents are real quantities, ψ_{II} does not oscillate and so does not represent a moving particle. However, the probability density $|\psi_{II}|^2$ is not zero, so there is a finite probability of finding a particle within the barrier. Such a particle may energy into region III or it may return to region I.

Transmission Probability

In order to calculate the transmission probability T , we have to apply appropriate boundary conditions to ψ_I , ψ_{II} and ψ_{III} . Figure 1.24 shows the wave functions in regions I, II and III. As discussed earlier, both ψ and its derivative $\frac{\partial\psi}{\partial x}$ must be continuous everywhere. With reference to Fig. 1.24, these conditions mean that for a perfect fit at each side of the barrier, the wave functions inside and outside must have the same value and the same slope. Hence at the L.H.S. of the barrier

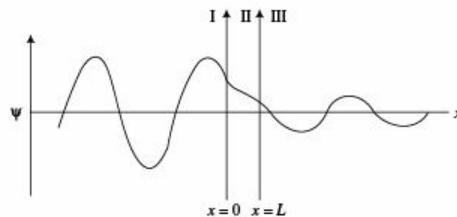


Fig. 1.24 At each wall of the barrier, the wave functions inside and outside it must match up perfectly, which means that they must have the same values and slopes there.

boundary conditions at $x = 0$
$$\left. \begin{aligned} \psi_I &= \psi_{II} \\ \frac{d\psi_I}{dx} &= \frac{d\psi_{II}}{dx} \end{aligned} \right\}_{x=0} \quad \dots(1.114)$$

and at right-hand side, i.e., boundary condition at $x = L$

$$\left. \begin{aligned} \psi_{II} &= \psi_{III} \\ \frac{d\psi_{II}}{dx} &= \frac{d\psi_{III}}{dx} \end{aligned} \right\}_{x=L} \quad \dots(1.115)$$

Now we substitute ψ_I , ψ_{II} and ψ_{III} from Eqs. (1.101), (1.102) and (1.107) into the above equations. This yields in the same order

$$A + B = C + D \quad \dots(1.116)$$

$$ik_1 A - ik_1 B = -k_2 C + k_2 D \quad \dots(1.117)$$

$$Ce^{-k_2 L} + De^{k_2 L} = Fe^{ik_1 L} \quad \dots(1.118)$$

$$-k_2 Ce^{-k_2 L} + k_2 De^{k_2 L} = ik_1 F e^{ik_1 L} \quad \dots(1.119)$$

Eqs. (1.116) to (1.119) may be solved for (A/F) to give

$$\frac{A}{F} = \left[\frac{1}{2} + \frac{i}{4} \left(\frac{k_2}{k_1} - \frac{k_1}{k_2} \right) \right] e^{i(k_1 + k_2)L} + \left[\frac{1}{2} - \frac{i}{4} \left(\frac{k_2}{k_1} - \frac{k_1}{k_2} \right) \right] e^{i(k_1 - k_2)L} \quad \dots(1.120)$$

Approximations

Let us assume that the potential barrier U is high relative to the energy E of the incident particles. If this is the case, then

$$\frac{k_2}{k_1} \gg \frac{k_1}{k_2} \text{ and } \frac{k_2}{k_1} - \frac{k_1}{k_2} \approx \frac{k_2}{k_1} \quad \dots(1.121)$$

Let us also assume that the barrier is wide enough for ψ_{II} to be severely weakened between $x = 0$ and $x = L$. This means that $k_2 L \gg 1$ and

$$e^{k_2 L} \gg e^{-k_2 L}$$

Hence, Eq. (1.120) can be approximated by

$$\frac{A}{F} = \left(\frac{1}{2} + \frac{i}{4} \frac{k_2}{k_1} \right) e^{i(k_1 + k_2)L} \quad \dots(1.122)$$

Computation of T

The complex conjugate of (A/F) , which we need to compute the transmission probability T , is found by replacing i by $-i$ wherever it occurs in (A/F) , we have

$$\left(\frac{A}{F} \right)^* = \left(\frac{1}{2} - \frac{ik_2}{4k_1} \right) e^{-ik_1 + k_2 L} \quad \dots(1.123)$$

Multiplying (A/F) and $(A/F)^*$, we get

$$\frac{AA^*}{FF^*} = \left(\frac{1}{4} + \frac{k_2^2}{16k_1^2} \right) e^{2k_2 L}$$

Here, $v_{III+} = v_{I+}$ so $\frac{v_{III+}}{v_{I+}} = 1$ in Eq. (1.109), which means that the transmission probability is

$$T = \frac{FF^* v_{III+}}{AA^* v_{I+}} = \left(\frac{AA^*}{FF^*} \right)^{-1} = \left[\frac{16}{4 + \left(\frac{k_2}{k_1} \right)^2} \right] e^{-2k_2 L} \quad \dots(1.124)$$

From the definitions of k_1 and k_2 (Eq. 1.112), we see that

$$\left(\frac{k_2}{k_1} \right)^2 = \frac{2m(U - E)/\hbar^2}{2mE/\hbar^2} = \frac{U}{E} - 1 \quad \dots(1.125)$$

This formula means that the quantity in brackets in Eq. (1.124) varies much less with E and U than does the exponential. The bracketed quantity, furthermore, always is of the order of magnitude of 1 in value. A reasonable approximation of the transmission probability is therefore,

$$T = e^{-2k_2 L} \quad \dots(1.126)$$

This is the approximate transmission probability expression showing that larger is L , smaller is the transmission probability. Thus, the transmission probability depends on two variables k_2 and L (the thickness of the barrier). In many cases, the product $k_2 L$ is very large, making the probability very small. Nevertheless, if the particle keeps striking the barrier, it will eventually pass through it as in the case of α -decay.

The concept of tunnelling has been successfully used to explain α -decay from radioactive nuclei, working of tunnel diode and electron tunnelling in scanning tunnelling microscope (STM).

Example 1.26

Consider a particle in the ground state in a rigid box of length L . (a) Find the probability density $|\psi|^2$. (b) Where is the particle most likely to be found? (c) What is the probability of finding the particle in the interval between $x = 0.50L$ to $x = 0.51L$? (d) What is it for the interval $[0.75L, 0.76L]$? (e) What would be the average result if the position of a particle in the ground state were measured many times?

Solution.

- (a) The probability density $|\psi(x)|^2$ for the ground state $n = 1$ is

$$|\psi(x)|^2 = \frac{2}{L} \sin^2 \left(\frac{\pi x}{L} \right)$$

which is plotted in Fig. 1.25.

- (b) The most probable value x , is the value of x for which $|\psi(x)|^2$ is maximum, which is clearly seen to be

$$x_{mp} = \frac{L}{2}$$

- (c) The probability of finding the particle in any small interval from x to $x + \Delta x$ is $|\psi(x)|^2 \Delta x$. (This is exact in the limit $\Delta x \rightarrow 0$ and

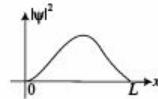


Fig. 1.25 Probability density $|\psi(x)|^2$ for a particle in the ground state in a rigid box.

therefore a good approximation for small interval Δx

$$\therefore P(0.50L \leq x \leq 0.51L) \approx |\Psi(0.50L)|^2 \Delta x = \frac{2}{L} \sin^2\left(\frac{\pi}{2}\right) \times 0.01 L \\ = 0.02 \text{ or } 2\%$$

$$(d) \text{ Similarly, } P(0.75 \leq x \leq 0.76L) \approx \frac{2}{L} \sin^2\left(\frac{3\pi}{4}\right) \times 0.01L = 0.01 \text{ or } 1\%$$

(e) The average result would be the expectation value given by

$$\langle x \rangle = \frac{2}{L} \int_0^L x \sin^2\left(\frac{\pi x}{L}\right) dx$$

which can be seen to be $\frac{x}{2}$ by evaluating the integral.

Example 1.27

Find the answers of the above questions for the first excited state.

Solution.

The probability density for $n = 2$ is given by

$$|\psi(x)|^2 = \frac{2}{L} \sin^2\left(\frac{2\pi x}{L}\right)$$

and it is seen from the plotted figure that

$$x_{mp} = \frac{4}{L} \text{ and } \frac{3L}{4}.$$

Also, $P(0.50L \leq x \leq 0.51L) \approx |\psi(0.50L)|^2 \Delta x = 0$, since $\psi(0.50L) = 0$ [Strictly speaking, the probability for this interval is not exactly zero since the probability density $|\psi(x)|^2$ is zero at only one point $0.50L$. In fact, the probability for this interval is very small compared to the probability for intervals of the same width elsewhere.]

$$P(0.75L \leq x \leq 0.76L) \approx |\psi(0.75L)|^2 \Delta x = \frac{2}{L} \sin^2\left(\frac{3\pi}{2}\right) \times 0.01L \\ = 0.02 \text{ or } 2\%$$

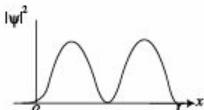


Fig. 1.26 Probability density $|\psi(x)|^2$ for $n = 2$.

Example 1.28

Suppose two identical conducting wires, lying along the x -axis are separated by an air gap of thickness $L = 1 \text{ nm}$ (\approx a few atomic diameters). An electron is moving inside either conductor has potential energy zero, whereas in the gap its potential energy $U > 0$. Thus, the gap is a kind of barrier discussed in the previous section. The electron approaches the barrier from the left with energy such that $U - E = 1 \text{ eV}$, i.e., the electron is 1 eV below the top of the barrier. Find the

probability of the electron emerging on the other side of the barrier? If the barrier is twice as wide, how the probability changes?

Solution.

The probability is given by $T = e^{-2k_2 L}$

$$k_2 = \frac{\sqrt{2m(U-E)}}{\hbar} = \frac{\sqrt{2mc^2(U-E)}}{\hbar c} = \frac{\sqrt{2 \times 5 \times 10^5 \text{eV} \times 1 \text{eV}}}{197 \text{eV} \cdot \text{nm}}$$

$$\approx 5.1 \text{ nm}^{-1}$$

$$\therefore T = e^{-2 \times 5.1 \text{ nm}^{-1} \times 1 \text{ nm}} = e^{-10.2} = 3.7 \times 10^{-5} \approx 0.04 \%$$

Thus, the probability is not very large. However, if large number of electrons approaches the barrier, a few of them will tunnel through.

For $L = 2 \text{ nm}$, $T = e^{-20.4} = 1.4 \times 10^{-9}$, much smaller than the previous one indicating the extreme sensitivity of the transmission probability to the gap width – a fact which is utilized in the scanning tunnelling microscope.

SHORT ANSWER TYPE QUESTIONS

1. Mention few limitations of classical mechanics.
2. What is a photon? Give its important characteristics.
3. Explain how interactions of photons with atoms can give (i) photoelectric effect, (ii) Compton effect, and (iii) pair-production.
4. What is photoelectric effect?
5. Define the terms: (i) stopping potential, (ii) threshold frequency and (iii) saturation current for photoelectric effect.
6. The photoelectric effect manifest particle or wave nature of light? Explain.
7. Write Einstein's photoelectric equation and discuss the existence of threshold frequency on the basis of this equation.
8. What is a photoelectric cell? Mention some applications of this cell.
9. What is Compton effect? Explain how it supports the photon nature of light.
10. Explain why Compton effect cannot be observed with visible light?
11. For what scattering angle the Compton shift is maximum?
12. What do you mean by pair-production? Mention the necessary conditions for it.
13. What do you mean by pair annihilation?
14. What do you mean by matter waves? Give their characteristic.
15. Obtain an expression for a de-Broglie wavelength of a moving particle.
16. Show that the de-Broglie wavelength of a moving electron is given by $\lambda = \frac{12.27}{\sqrt{V}} \text{ Å}$, when the electron is accelerated through a potential difference V .
17. What do you mean by a wave packet?

18. Differentiate between group and phase velocity.
19. State Heisenberg's uncertainty principle and use this to show the non-existence of electrons inside the nucleus.
20. What do you mean by wave function? Mention the necessary characteristics of a wave function.
21. What do you mean by bra-ket notation in quantum mechanics? Explain.
22. What is an operator? Discuss write operators for energy and momentum.
23. Energy quantization appears as a natural element in quantum mechanics. Elaborate.
24. What do you mean by eigenvalues and eigenfunctions?
25. Differentiate between degenerate and non-degenerate eigenfunctions.
26. What do you mean by zero point energy? Explain.
27. Existence of zero point energy is consistent with uncertainty principle. Explain.
28. What do you mean by tunnelling? Explain.
29. On what factors the transmission probability depends. Name the devices where tunnelling effect is utilized.

LONG ANSWER TYPE QUESTIONS

1. Obtain Planck's radiation law in terms of the frequency and wavelength. Show that Wien's displacement law and Rayleigh-Jean's law are special cases of Planck's law.
2. What is photoelectric effect? Explain. Discuss the effect of (i) intensity and (ii) frequency of incident radiation.
3. Discuss Einstein's photoelectric relation to show that the laws of photoelectric emission can be explained using this relation.
4. Explain Compton effect and its physical significance. Obtain an expression for the Compton shift.
5. What do you mean by matter waves? Discuss Davisson and Germer experiment to show the existence of matter waves.
6. A moving material particle is equivalent to a wave packet rather than a single wave. Explain.
7. State and explain Heisenberg uncertainty principle. Discuss its applications.
8. What do you mean by expectation value of a physical quantity? Explain why the expectation value of momentum cannot be calculated using this method?
9. Obtain time dependent form of Schrödinger equation. Use this equation to obtain time independent form of Schrödinger equation.
10. Write the Schrödinger equation for a particle confined in a box with potential energy $U = 0$ within the box. Solve this equation to obtain eigenvalues and eigenfunctions. Represent first three eigenfunctions and corresponding probability densities.
11. Write Schrödinger equation for a simple harmonic oscillator and discuss its solution. Represent the first four eigenfunctions and corresponding probability densities and discuss its important features.
12. Discuss tunnel effect and obtain the expression for transmission probability for a rectangular barrier.

NUMERICALS

1. Calculate the stopping potential for the photoelectrons emitted by a gold cathode if the wavelength of the incident radiation is 2×10^{-7} m. Work function of gold is 4.80 eV. [Ans. 1.41V]

2. In an experiment of Compton scattering the incident radiation has wavelength 2\AA , while wavelength of radiation scattered through 180° is 2.048\AA . Calculate (i) the wavelength of the scattered radiation if they are viewed at an angle of 60° to the direction of incidence, (ii) the energy of the recoil electron which scatters radiation through 60° .

[Ans. (i) 2.012\AA , (ii) $5.91 \times 10^{-18}\text{ J}$]

3. Calculate the de-Broglie wavelength associated with a proton moving with a velocity equal to $\frac{1}{20}$ th of the velocity of light.

[Ans. $2.634 \times 10^{-14}\text{ m}$]

4. Calculate the de-Broglie wavelength of a neutron of energy 28.8 eV . ($m_N = 1.67 \times 10^{-27}\text{ kg}$)

[Ans. 0.05\AA]

5. Through what voltage an electron needs to be accelerated so that its wavelength is 0.50\AA ?

[Ans. 602 V]

6. Calculate the de-Broglie wavelengths of a 2 eV photon and an electron with kinetic energy of 2 eV .

[Ans. 6200\AA , 8.6\AA]

7. The average time that an atom retains excess excitation energy before re-emitting it in the form of electromagnetic radiations is 10^{-8} sec . Calculate uncertainty in energy of the emitted radiation.

[Ans. $6.56 \times 10^{-8}\text{ eV}$]

8. A proton is confined in an one-dimensional box of length 10^{-8} m . Calculate the uncertainty in its velocity.

[Ans. $2.3 \times 10^{-7}\text{ m/s}$]

Unit II

2

Statistical Mechanics ---

2.1 INTRODUCTION

Statistical mechanics is the mechanics of systems with a very large number of constituent particles, e.g., a litre of air with nearly 10^{22} molecules of N₂ and O₂, a cup of water with some 10^{25} water molecules, a metre of copper wire with 10^{23} conduction electrons. The approach here is to take a number of measurements of an observable and then to evaluate its mean value or the most probable value. Statistical mechanics makes no attempt to obtain a complete solution for complex systems. Instead, it makes predictions about certain average properties by studying the motion of a large number of identical systems, ignoring the detailed motion of individual particles.

There are classical and quantum versions of statistical mechanics. We first discuss the basic rules of classical statistical mechanics and then use them to derive Maxwell-Boltzmann's distribution law and equipartition theorem. Later we will discuss quantum statistics of Bose-Einstein and Fermi-Dirac and their applications.

2.2 MICROSCOPIC AND MACROSCOPIC SYSTEMS

Let us consider a gas confined in a container. It is the collection of very large number of molecules or atoms of the order of 10^{20} or higher. A gas molecule or atom in a container is a *microscopic system* whereas the gas in the container is a *macroscopic system*. The detailed specification of the microscopic state or microstate of any system of N particles involves specifying the positions and velocities (in classical approach) of all the N particles and specifying wave functions of all N particles (in quantum approach). The number of particles N is so large that to write the microstate of a mole of a gas it will require more than India's total supply of paper, making it practically impossible. Fortunately, in statistical mechanics it is not attempted either. Rather, we want to know just a small number of **macroscopic variables** that are statistical averages over the many microscopic variables. For example, to specify the *macroscopic state* or *macrostate*, of a gas confined in a container, we need only to give the volume V, the number of atoms N, the total energy E and the temperature T.

In statistical mechanics, we study the relation between the observable macrostates and underlying (unobservable) microstates. The relations connecting microscopic and macroscopic properties help to understand the microscopic properties of the system. In the case of a gas confined in a container, kinetic theory gives the relationship between the microscopic motion of the molecules and the macroscopic variables, e.g., pressure and temperature.

2.3 PHASE SPACE

As mentioned above, in classical statistical mechanics, the state of a system can be completely defined by specifying the positions and velocities of all its constituent particles. A system having f degrees of freedom can be specified by means of f generalized coordinates $q_1, q_2, q_3, \dots, q_f$ and f generalized momenta p_1, p_2, \dots, p_f . The conceptual space having $2f$ dimensions is called the *phase space* which is supposed to have $2f$ rectangular axes, one for each of the generalized coordinates and one each for the corresponding momenta. The instantaneous state of a system having f degrees of freedom can be specified by a point, called *phase point* in the $2f$ dimensional space.

Ensemble. Collection of a large number of identical systems is generally referred to as an *ensemble of systems*. All the members of the ensemble must have the same structure, i.e., they must be identical w.r.t. their shape, size, number of molecules, total energy, etc. However, the systems can differ in the coordinates and velocities of the molecules. An ensemble is described in terms of a quantity called the *density of distribution* (D) with which the representative points are distributed in the phase space. When the ensemble contains a large number of systems, the density of points in the phase space can be treated as a continuous function. The density of distribution D of the phase points is expected to be a function of the $2f$ coordinates $q_1, q_2, \dots, q_f; p_1, p_2, \dots, p_f$; and time t . That is,

$$D = D(q, p, t) \quad \dots(2.1)$$

where q stands for q_1, q_2, \dots, q_f and p stands for p_1, p_2, \dots, p_f . The quantity $D(dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f)$ gives the number of systems dn lying in the volume element $d\Omega$ around the point of interest. Then

$$D(q, p, t) = \frac{dn}{d\Omega} \quad \dots(2.2)$$

where $d\Omega = dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f$.

Liouville's Theorem: According to the Liouville's theorem, '*the rate of change of phase density in the immediate vicinity of any given phase point as it moves through phase space is zero*'.

That is,

$$\frac{dD}{dt} = 0 \quad \dots(2.3)$$

This implies that the volume in phase space occupied by a definite number of phase points does not change with time. In the particular case, when the probabilities of finding the phase points in the various regions of the phase space are independent of time,

$$\left(\frac{\partial D}{\partial t} \right)_{q,p} = 0 \quad \dots(2.4)$$

Such an ensemble is said to be in *statistical equilibrium*.

2.4 EQUAL PROBABILITY HYPOTHESIS

The entire statistical mechanics is based on a hypothesis often called *postulate of equal probability* which states that *the probability of finding the phase point in any one region of phase space is identical with that for any other region of equal volume*. It can be shown that this hypothesis is in conformity with the Liouville's theorem. Also, the predictions based this postulate are in agreement with experimental observations justifying the postulate.

While dealing with molecules, it is convenient to consider the phase space of the molecules to be divided into a number of elementary regions, called cells, each of volume $dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f$.

2.5 STATISTICAL EQUILIBRIUM

Consider a thermodynamical isolated system consisting of N particles. The energy states available to the particles are E_1, E_2, E_3 , etc. These energy states may be quantized or may be continuous and are due to vibrational and rotational energy of the particles.

Suppose, that at any given instant of time, n_1 particles are in state of energy E_1 , n_2 particles with energy state E_2 , and so on.

The total number of particles in the system is,

$$N = n_1 + n_2 + n_3 + \dots = \sum_i n_i, i = 1, 2, 3 \dots \quad \dots(2.5)$$

The total energy of the system

$$\begin{aligned} E &= n_1 E_1 + n_2 E_2 + n_3 E_3 + \dots \\ E &= \sum_i n_i E_i \end{aligned} \quad \dots(2.6)$$

Equation (2.6) refers to the total energy of a system in which the particles are non-interacting. Here, the energy of each particle depends only on the coordinates of particle in the system for an isolated system, the total energy E is constant.

$$\text{For isolated system, } E = \sum_i n_i E_i = \text{constant}$$

Consider a gas having N molecules at a certain temperature and pressure. Its volume, temperature and pressure are kept constant, i.e., the system is isolated. The total energy of this system remains constant. But the molecules of the gas collide with each other and also with the walls of the container. Consequently, the number of molecules change from one energy state to the other energy state. It means that the values of n_1, n_2, n_3 , etc. continuously change. It can be reasonably assumed that for each microscopic state of a system of particles, there is a particular most favoured distribution. When this distribution for the particles is reached, the system attains statistical equilibrium. For an isolated system, the values of n_1, n_2, n_3 , etc. vary only near the values corresponding to the most probable distribution. Hence, the basic problem in statistical mechanics is to obtain the most probable distribution law for a given composition of the system.

In practice, three most probable distribution laws are used. They are:

- (i) Maxwell-Boltzmann Distribution Law
- (ii) Fermi-Dirac Distribution Law
- (iii) Bose-Einstein Distribution Law

2.6 MAXWELL-BOLTZMANN DISTRIBUTION LAW

Consider a system having n similar but distinguishable particles (or molecules) in a container with its total energy E . Let the individual molecules are in different phase cells (or energy levels). The probability W of finding the system in the state in which n_1 molecules are in the state with energy E_1 , n_2 with the energy E_2 , and so on, is given by

$$W = \frac{n!}{n_1! n_2! \dots n_i!} \times \text{constant} \quad \dots(2.7)$$

Taking logarithm of Eq. (2.7), we get

$$\ln W = \ln n! - \sum_i \ln n_i! + \text{constant} \quad \dots(2.8)$$

In statistical systems, n and n_i are very large and for large values of n , Stirling's approximation gives

$$\ln n! = n \ln n - n \quad \dots(2.9)$$

Using this, Eq. (2.8) becomes

$$\begin{aligned} \ln W &= n \ln n - n - \sum_i (n_i \ln n_i - n_i) + \text{constant} \\ &= n \ln n - \sum_i n_i \ln n_i + \text{constant} \end{aligned} \quad \dots(2.10)$$

For the probability to be maximum, which corresponds to the most probable distribution of molecules among the energy levels

$$d(\ln W) = 0 \quad \dots(2.11)$$

Substituting $\ln W$, we get

$$\sum_i (\ln n_i + 1) dn_i = 0 \quad \dots(2.12)$$

The condition for maximum probability given by Eq. (2.12) is subject to the following two conditions.

- (i) The total number of molecules n is a constant. That is,

$$\sum_i n_i = n = \text{constant}$$

$$\text{or} \quad \sum_i dn_i = 0 \quad \dots(2.13)$$

- (ii) The total energy E is a constant (although the energy of the molecule in the i th state E_i may vary). That is,

$$\sum_i E_i n_i = E = \text{constant}$$

$$\sum_i E_i dn_i = 0 \quad \dots(2.14)$$

Equations (2.12), (2.13) and (2.14) are the conditions to be satisfied simultaneously for the system to have the maximum probability. The most probable distribution of the molecules among various energy levels can now be obtained by Lagrange's method of undetermined multipliers. Multiplying Eq. (2.13) by α' and Eq. (2.14) by β and then adding to Eq. (2.12), we get

$$\sum_i (\ln n_i + 1 + \alpha' + \beta E_i) dn_i = 0 \quad \dots(2.15)$$

The constant $(1 + \alpha')$ can be replaced by another constant α and as the variations dn_i 's are independent, Eq. (2.15) will be satisfied only if each term in the summation is zero. That is,

$$\ln n_i + \alpha + \beta E_i = 0$$

or

$$\ln n_i = -(\alpha + \beta E_i)$$

or

$$n_i = \frac{1}{\exp(\alpha + \beta E_i)} \quad \dots(2.16)$$

This is Maxwell-Boltzmann distribution law and it gives the most probable distribution of molecules among the various possible individual energy values.

It can be seen that the constant $\beta = \frac{1}{kT}$, where $k (= 1.3805 \times 10^{-23} \text{ J/K})$ is the Boltzmann constant and T is temperature. Here, kT has the unit of energy (Joule). Therefore, the Maxwell Boltzmann distribution can also be expressed as

$$n_i = \frac{1}{\exp\left(\alpha + \frac{E_i}{kT}\right)} \quad \dots(2.17)$$

The constant α depends on the volume and the temperature of the gas.

2.7 MAXWELL'S SPEED DISTRIBUTION LAW

The velocity distribution law was derived by Maxwell on the basis of kinetic theory of gases. Suppose n molecules of a gas is contained in a container of volume V . From Maxwell-Boltzmann distribution law, we have

$$n_i = e^{-\alpha} e^{-\beta E_i} \quad \dots(2.18)$$

For monoatomic gases, the value of $e^{-\alpha}$ is given by

$$e^{-\alpha} = \frac{n dx dy dz dp_x dp_y dp_z}{V (2\pi m k T)^{3/2}} \quad \dots(2.19)$$

Substituting the values of $e^{-\alpha}$ and β in Eq. (2.18), replacing n_i – the number of molecules in the i^{th} state by dn and omitting the subscript i from E_i , we get

$$dn = \frac{n \exp\left(-\frac{E}{kT}\right) dx dy dz dp_x dp_y dp_z}{V (2\pi m k T)^{3/2}} \quad \dots(2.20)$$

Integrating over the coordinates x, y, z (so that $\int \int \int dx dy dz = V$), yields

$$dn = \frac{n \exp(-E/kT)}{(2\pi mkT)^{3/2}} dp_x dp_y dp_z \quad \dots(2.21)$$

Here, dn represents the number of molecules in the container with momentum components in the interval dp_x, dp_y and dp_z . The translational energy E of a molecule having mass m is given by

$$E = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) \quad \dots(2.22)$$

while its momentum components can be represented as

$$p_x = mv_x, \quad p_y = mv_y, \quad p_z = mv_z \quad \dots(2.23)$$

$$\text{Therefore, } dp_x dp_y dp_z = m^3 dv_x dv_y dv_z \quad \dots(2.24)$$

Substituting these in Eq. (2.21), we have

$$dn = n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left[\frac{-m(v_x^2 + v_y^2 + v_z^2)}{2kT} \right] dv_x dv_y dv_z \quad \dots(2.25)$$

This is *Maxwell's distribution law* for the number of molecules with velocity components in the interval dv_x, dv_y and dv_z . Writing the velocity components v_x, v_y and v_z in terms of the speed of a molecule v and changing over to polar coordinates, so that

$$v^2 = v_x^2 + v_y^2 + v_z^2 \quad \text{and} \quad dv_x dv_y dv_z = v^2 \sin \theta d\theta d\phi dv.$$

Substituting these values in Eq. (2.25), we get

$$dn = n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(\frac{-mv^2}{2kT} \right)^{3/2} v^2 \sin \theta d\theta d\phi dv \quad \dots(2.26)$$

This equation gives the number of molecules with speeds between v and $v + dv$ in a direction lying within the angular range θ to $\theta + d\theta$ and ϕ to $\phi + d\phi$. Integration over the angular part gives the number of molecules with velocity between v and $v + dv$

$$\begin{aligned} dn &= n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(\frac{-mv^2}{2kT} \right) v^2 dv \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \\ &= 4\pi n \left(\frac{m}{2\pi kT} \right)^{3/2} v^2 \exp \left(\frac{-mv^2}{2kT} \right) dv \end{aligned} \quad \dots(2.27)$$

Using $F(v)dv$ to denote the number of molecules having speeds between v and $v + dv$, we have

$$F(v)dv = 4\pi n \left(\frac{m}{2\pi kT} \right)^{3/2} v^2 \exp \left(\frac{-mv^2}{2kT} \right) \dots(2.28)$$

The function $F(v)dv$ is called the *Maxwell speed distribution function* derived by Maxwell before development of quantum mechanics or quantum statistics. The function $F(v)$ is represented in Fig. 2.1.

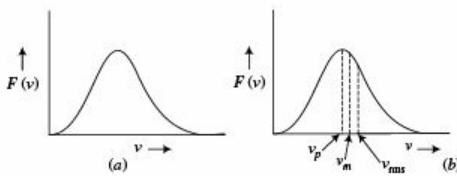


Fig. 2.1 (a) The Maxwell speed distribution (b) most probable speed (v_p), mean speed (v_m) and rms speed (v_{rms}) marked on speed distribution.

The Maxwell speed distribution has been experimentally verified. The Maxwell speed distribution for O_2 and comparatively lighter H_2 molecules are shown in Fig. 2.2. In both the cases there is tail out at higher speeds, but few molecules have speeds higher than about three times the most probable speed, which would correspond to an energy nine times the most probable energy. This indicates why there is almost no hydrogen in the earth's atmosphere. For hydrogen, the tail of the speed distribution extends beyond the earth's escape speed (≈ 10 km/s), so that any hydrogen in the earth's atmosphere eventually leaks into the space. Smaller astronomical bodies such as moon and asteroids have such low escape velocities that all gases on their surfaces escape to space.

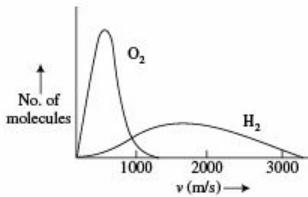


Fig. 2.2

Since the kinetic energy $E = \frac{1}{2}mv^2$, we can write

$$v^2 dv = \frac{(2E)^{1/2} dE}{m^{3/2}} \quad \dots(2.29)$$

Substituting these values of E and $v^2 dv$, Eq. (2.27) can be expressed as

$$dn = \frac{2\pi n}{(\pi kT)^{3/2}} \exp(-E/kT) E^{1/2} dE \quad \dots(2.30)$$

which gives the number of molecules (or atoms) having energy between E and $E + dE$. The *most probable speed* v_p corresponds to the peak of $F(v)$ versus curve (Fig. 2.1(b)). Using

Eq. (2.28), for $F(v)$ to be maximum

$$\frac{d}{d}F(v) = 4\pi n \left(\frac{m}{2\pi kT} \right)^{3/2} \left(2v_p - v_p^3 \frac{m}{kT} \right) \exp \left(\frac{-mv^2}{2kT} \right) = 0.$$

or $2v_p - v_p^3 \frac{m}{kT} = 0$

or $v_p = \sqrt{\frac{2kT}{m}}$... (2.31)

The *mean speed* v_m is calculated as

$$v_m = \frac{1}{n} \int_0^\infty v F(v) dv = \frac{1}{n} 4\pi n \left(\frac{m}{2\pi kT} \right)^{3/2} \int_0^\infty v^3 \exp \left(\frac{-mv^2}{2kT} \right) dv$$

Putting the value of the integral

$$v_m = 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} \cdot \frac{1}{2} \left(\frac{2kT}{m} \right)^2$$

or $v_m = \sqrt{\frac{8kT}{\pi m}}$... (2.32)

The *root mean square speed* v_{rms} is given by

$$\begin{aligned} v_{rms}^2 &= \bar{v}^2 = \frac{1}{n} \int_0^\infty v^2 F(v) dv \\ &= 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} \int_0^\infty v^4 \exp \left(\frac{-mv^2}{2kT} \right) dv \end{aligned}$$

Putting the value of the integral

$$v_{rms}^2 = 4\pi \left(\frac{m}{2\pi kT} \right)^2 \cdot \frac{3\sqrt{\pi}}{8} \cdot \left(\frac{2kT}{m} \right)^{5/2}$$

or $v_{rms} = \sqrt{\frac{3kT}{m}}$... (2.33)

The values of v_p , v_m and v_{rms} are represented in Fig. 2.1 (b).

2.8 LAW OF EQUIPARTITION OF ENERGY

The expression for the number of molecules having coordinates in the interval dx , dy and dz and momentum components in the interval dp_x , dp_y and dp_z , given by Eq. (2.21) enables us to evaluate mean values in statistical mechanics. If A is any quantity which is a function of the three coordinates

and the three momenta, the mean value of A , expressed as \bar{A} , is given by

$$\bar{A} = \frac{\int \dots \int A dn}{\int \dots \int dn} = \frac{\int \dots \int A e^{-E/kT} dx dy dz dp_x dp_y dp_z}{\int \dots \int e^{-E/kT} dx dy dz dp_x dp_y dp_z} \quad \dots(2.34)$$

Accordingly, the average energy of molecules in a system is given by

$$\bar{E} = \frac{\int \dots \int E e^{-E/kT} dx dy dz dp_x dp_y dp_z}{\int \dots \int e^{-E/kT} dx dy dz dp_x dp_y dp_z} \quad \dots(2.35)$$

Here, it is being assumed that the energy depends on all the variables. Let us assume that the energy $E(p_x)$ associated with momentum p_x is separable from the rest of the energy, say E' , which is dependent on all the other variables. That is,

$$E = E(p_x) + E' \quad \dots(2.36)$$

The average value of $E(p_x)$ is

$$\bar{E}(p_x) = \frac{\int \dots \int E(p_x) e^{-E(p_x)/kT} e^{-E'/kT} dx dy dz dp_x dp_y dp_z}{\int \dots \int e^{-E(p_x)/kT} e^{-E'/kT} dx dy dz dp_x dp_y dp_z} \quad \dots(2.37)$$

Since $E(p_x)$ depends only on the variable p_x , therefore

$$\bar{E}(p_x) = \frac{\int E(p_x) e^{-E(p_x)/kT} dp_x}{\int e^{-E(p_x)/kT} dp_x} \quad \dots(2.38)$$

Expressing the energy as a quadratic function of momenta

$$E(p_x) = a p_x^2 \quad \dots(2.39)$$

where a is a constant. Substituting this in $E(p_x)$, we get

$$\bar{E}(p_x) = \frac{\int a p_x^2 e^{-ap_x^2/kT} dp_x}{\int e^{-ap_x^2/kT} dp_x} \quad \dots(2.40)$$

The solution of the integrals on the R.H.S yields $\frac{1}{2}kT$. Therefore,

$$\bar{E}(p_x) = \frac{1}{2}kT \quad \dots(2.41)$$

Thus, the average (kinetic) energy associated with a variable, which contributes a quadratic term to the total energy, is $\frac{1}{2}kT$ per molecule. This is referred to as the law of equipartition of

energy. This is also expressed in terms of the degree of freedom, wherein it is stated that the average kinetic energy associated with each degree of freedom is $\frac{1}{2}kT$. The total number of independent quantities which must be known for describing the state of a motion of a body are called its degrees of freedom.

2.9 QUANTUM STATISTICS

In classical statistics, the molecules though similar are treated as distinguishable from one another. This implies that the molecular interchange give a new microscopic state. In quantum statistics, however, particles are described by wave functions and identical particles cannot be distinguished from one another. This property of *indistinguishability* makes quantum statistics different from classical statistics. The distribution function for the indistinguishable particles is different from the classical distinguishable particles.

Symmetric wave function (ψ_s): A wave function is *symmetric*, if the interchange of any pair of particles among its argument leaves the wave function unchanged.

Anti-symmetric wave function (ψ_A): A wave function is antisymmetric, if the interchange of any pair of particles among its arguments changes the sign of the wave function.

If P is an exchange operator, then

$$P\psi_s(1, 2) = \psi_s(2, 1) \quad \dots(2.42)$$

and $P\psi_A(1, 2) = -\psi_A(2, 1) \quad \dots(2.43)$

The symmetry property of the wave function is related with the spin of the particle.

(i) The identical particles having an *integral spin quantum number* are described by symmetric wave functions, i.e.,

$$P\psi_s(1, 2, 3, \dots, r, s, \dots, n) = \psi_s(1, 2, 3, \dots, s, r, \dots, n)$$

Such particles described by symmetric wave functions and having integral spins, do not obey Pauli's principle and are called *bosons* and obey *Bose-Einstein statistics*. Photons and π -mesons are bosons. Atoms and molecules composed of even number of fermions (half integral spins) are also bosons since their total spin will be zero or an integer.

(ii) The identical particles having half odd integral spin quantum number are described by anti-symmetrical wave functions, i.e.,

$$P\psi_A(1, 2, 3, \dots, r, s, \dots, n) = -\psi_A(1, 2, \dots, s, r, \dots, n)$$

Such particles described by antisymmetric wave functions having half odd integral spins obey Pauli's principle and are called *fermions* and obey *Fermi-Dirac statistics*. Electron, proton, neutron, muon, etc. are fermions.

The symmetry restrictions on the wave functions are not there if the particles are non-interacting and distinguishable. In such cases Maxwell-Boltzmann statistics are applicable.

2.10 BOSE-EINSTEIN STATISTICS

This statistics is followed by identical, indistinguishable particles of integral spin (bosons) described by symmetrical wave functions. It was developed by S.N. Bose for photons and generalized by Einstein.

Let us consider a system consisting of n indistinguishable particles. Let these particles be divided into different quantum groups such that n_1 particles are in level 1 having energy E_1 , n_2 are in level 2 having energy E_2 , n_3 are in level 3 having energy E_3 , and so on. Since the particles are indistinguishable there is only one way in which subdivisions can be made. The total number of particles is a constant, i.e.,

$$\sum_i n_i = n = \text{Constant} \quad \dots(2.44)$$

The number n_i is very large and therefore can be treated as a continuous variable. Hence,

$$\sum_i dn_i = dn = 0 \quad \dots(2.45)$$

The constancy of the total energy gives

$$\sum_i E_i n_i = E = \text{Constant} \quad \dots(2.46)$$

or

$$\sum_i E_i dn_i = 0 \quad \dots(2.46(a))$$

Energy levels are often degenerate. Let g_i is the degeneracy of the i th level (for non-degenerate case $g_i = 1$). This means n_i particles are to be distributed to the g_i wave functions. In Bose-Einstein statistics there is no restriction on the number of particles associated with a particular wave function. The number of ways of distributing n_i particles in the i th level which is g_i fold degenerate is equivalent to the distribution of n_i indistinguishable particle in a box divided into g_i compartments without any restriction on the number in each compartment. The g_i compartments of the box can be obtained by $(g_i - 1)$ partitions. The total number of permutations of the n_i particles and the $(g_i - 1)$ partitions is $(n_i + g_i - 1)!$ Since the n_i particles are indistinguishable, total permutations among themselves, which are $n_i!$, do not give a new arrangement. Hence, the total number has to be divided by $n_i!$ In the same way, the permutation of $(g_i - 1)$ partitions do not alter the fact that there are still g_i sections. Therefore, the division of the total number by $(g_i - 1)!$ is also needed. Thus, the number of distinct ways of distributing n_i indistinguishable particles among the g_i wave function (or compartments) is

$$\frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!} \quad \dots(2.47)$$

Similarly, we can obtain similar expressions for other quantum states. Thus, the total number of ways G for the whole system can be expressed as

$$\begin{aligned} G &= \frac{(n_1 + g_1 - 1)! (n_2 + g_2 - 1)! \dots (n_i + g_i - 1)!}{n_1!(g_1 - 1)! n_2!(g_2 - 1)! \dots n_i!(g_i - 1)!} \\ &= \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!} \quad \dots(2.48) \end{aligned}$$

where the symbol \prod_i indicates the product of a series of similar terms.

In accordance with the postulate of equal probability of states, the probability W of the system for occurring with the specified distribution is proportional to the total number of eigenstates,

Hence,

$$W = \prod_i \frac{(n_i + g_i - 1)!}{n_i! (g_i - 1)!} \times \text{const.} \quad \dots(2.49)$$

Taking logarithm of Eq. (2.49)

$$\ln W = \sum_i [\ln(n_i + g_i - 1)! - \ln n_i! - \ln(g_i - 1)!] + \text{const.} \quad \dots(2.50)$$

Since n_i and g_i are very large numbers using Stirling's approximation ($\ln n_i! = n \ln n - n$), we have

$$\ln(g_i - 1)! = (g_i - 1) \ln(g_i - 1) - (g_i - 1) = g_i \ln g_i - g_i$$

$$\text{and} \quad \ln(n_i + g_i - 1)! \approx (n_i + g_i) \ln(n_i + g_i) - (n_i + g_i)$$

Substituting these in Eq. (2.50), we get

$$\ln W = \sum_i [(n_i + g_i) \ln(n_i + g_i) - n_i \ln n_i - g_i \ln g_i] + \text{const.} \quad \dots(2.51)$$

The condition for the probability to be maximum is

$$d(\ln W) = 0 \quad \dots(2.52)$$

Using this condition and remembering that g_i is not a variable while n_i is a continuous variable, Eq. (2.51) gives

$$d(\ln W) = \sum_i [\ln(n_i + g_i) + (n_i + g_i) \frac{1}{(n_i + g_i)} - \ln n_i - n_i \frac{1}{n_i}] dn_i = 0$$

$$\text{or} \quad \sum_i [\ln(n_i + g_i) - \ln n_i] dn_i = 0 \quad \dots(2.52a)$$

This condition for maximum probability is subject to the conditions given by Eqs. (2.45) and (2.46). Hence, the *most probable distribution* can be obtained by the Lagrange's method of undetermined multipliers. For this, multiplying Eq. (2.45) by $-\alpha$ and Eq. (2.46) by $-\beta$ and then adding to Eq. (2.52), we have

$$\sum_i [\ln \frac{n_i + g_i}{\ln n_i} - \alpha - \beta E_i] dn_i = 0 \quad \dots(2.53)$$

Since the variations dn_i are independent of each other, Eq. (2.53) will be satisfied only if the coefficient of each term in the summation is zero. That is,

$$\ln \frac{n_i + g_i}{n_i} - \alpha - \beta E_i = 0$$

$$\text{or} \quad \ln \frac{n_i + g_i}{n_i} = \alpha + \beta E_i \quad \dots(2.54)$$

Taking exponential on both sides, we get

$$\frac{n_i + g_i}{n_i} = e^{\alpha + \beta E_i}$$

or $\frac{g_i}{n_i} = e^{\alpha + \beta E_i} - 1$

or $n_i = \frac{g_i}{\exp(\alpha + \beta E_i) - 1} \quad \dots(2.55)$

This equation represents the most probable distribution of the particles among various energy levels for a system obeying Bose-Einstein statistics and is, therefore, known as *Bose-Einstein's distribution law*.

Bose-Einstein condensation. In 1924, Einstein predicted that a gas of identical bosons would undergo a new kind of phase transition if cooled to a low enough temperature. He showed that *below a certain critical temperature, most of the bosons in a gas will assemble into the ground state*. This low-temperature collection of ground-state bosons is called the *Bose-Einstein condensate*.

More than 70 years after Einstein's prediction, Bose-Einstein condensation was indeed achieved in June 1995 by Carl Wieman and Eric Cornell at the University of Colorado at Boulder in a gas of rubidium 87 atoms. They were awarded Nobel Prize in physics in 2001 for their experimental work on Bose-Einstein condensation. They achieved this at a temperature of 10^{-7} K in a sample of 10^{17} atoms evaporated down to 2000 atoms in the laser assisted cooling process.

2.11 FERMI-DIRAC STATISTICS

This statistics is followed by identical and indistinguishable particles of half-integral spin (fermions) described by anti-symmetrical wave functions.

Consider a system consisting of n indistinguishable particles. Let these n particles be divided into different quantum groups such that n_1 particles are in level 1 having energy E_1 , n_2 are in level 2 having energy E_2 , n_3 are in level 3 having energy E_3 , and so on. In this distribution following hold.

- (i) The particles are indistinguishable and therefore there is no distinction in which n_i particles are chosen.
- (ii) The particles obey Pauli's exclusion principle and therefore each sub-level or cell can have 0 or 1 particle. In other words, there can be only one particle in each of the g_i wave functions. Therefore, g_i must be greater than or equal to n_i because there must be at least one wave function for every particle in the group.

The total number of particles n is a constant and the total energy E is also a constant, as given by Eqs. (2.44) and (2.46) respectively. Due to Pauli's principle, no cell can accommodate more than one particle, therefore, among g_i cells n_i is occupied and $(g_i - n_i)$ cells are empty. The possible number of such a distribution is given by $g_i!$ corresponding to the permutations of the g_i cells. As the particles are indistinguishable, the occupied n_i cells are similar to each other and hence $n_i!$ permutations of occupied cells do not give distinguishable arrangement. Also $(g_i - n_i)!$ permutations of empty cells among themselves give rise to indistinguishable arrangements. Hence, the number of distinguishable arrangements of n_i particles in g_i cells is given by

$$g_i C_{n_i} = \frac{g_i!}{n_i!(g_i - n_i)!} \quad \dots(2.56)$$

Similarly, we can obtain expressions for other quantum states. The total number of eigenstates G for the whole system is

$$G = \frac{g_1!}{n_1!(g_1 - n_1)!} \cdot \frac{g_2!}{n_2!(g_2 - n_2)!} \cdots \cdots \frac{g_i!}{n_i!(g_i - n_i)!} \cdots \cdots \\ = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!} \quad \dots(2.57)$$

In accordance with the postulate of equal probability of states, the probability W of the system with the specified distribution is proportional to the total number of eigenstates. That is,

$$W = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!} \times \text{Constant} \quad \dots(2.58)$$

Taking logarithm on both sides,

$$\ln W = \sum_i [\ln g_i! - \ln n_i! - \ln(g_i - n_i)!] + \text{const.} \quad \dots(2.59)$$

As n_i , g_i and $(g_i - n_i)$ are large numbers, using Stirling approximation, we have

$$\ln W = \sum_i g_i \ln g_i - n_i \ln n_i - (g_i - n_i) \ln(g_i - n_i) + \text{const.} \quad \dots(2.60)$$

For the probability to be maximum, $d(\ln W) = 0$, and remembering that g_i is not variable while n_i is, we have

$$\sum_i \left(\ln \frac{g_i - n_i}{n_i} \right) dn_i = 0 \quad \dots(2.61)$$

The conditions given by Eqs. (2.44) and (2.46a) hold in this case also. To solve by the method of Lagrange's multipliers, multiplying Eqs. (2.44) and (2.46a) by constants $-\alpha$ and $-\beta$ respectively and then adding to Eq. (2.61), we get

$$\sum_i \left(\ln \frac{g_i - n_i}{n_i} - \alpha - \beta E_i \right) dn_i = 0 \quad \dots(2.62)$$

As the dn_i 's are independent

$$\ln \frac{g_i - n_i}{n_i} - \alpha - \beta E_i = 0$$

$$\text{or} \quad \ln \frac{g_i - n_i}{n_i} = \alpha + \beta E_i$$

Taking exponents on both sides, we get

$$\frac{g_i - n_i}{n_i} = \exp(\alpha + \beta E_i)$$

$$\text{or} \quad \frac{g_i}{n_i} = \exp(\alpha + \beta E_i) + 1$$

$$\text{or} \quad n_i = \frac{g_i}{\exp(\alpha + \beta E_i) + 1} \quad \dots(2.63)$$

This is the most probable distribution of n indistinguishable particles among the various energy levels obeying Fermi-Dirac statistics and is therefore known as *Fermi-Dirac distribution law*.

2.12 FERMI DISTRIBUTION FUNCTION–FERMI ENERGY

The constants α in the three distribution laws (*MB*, *BE* and *FD*) are different while the constant β $\left(= \frac{1}{kT} \right)$, where k is Boltzmann constant and T is the temperature in Kelvin is the same for all distributions. The distribution for the Fermi-Dirac statistics, Eq. (2.63) has the form

$$n(E) = g(E) F_{FD}(E) \quad \dots(2.64)$$

where $n(E)$ is a distribution and $n(E)dE$ represents the number of particles with energies between E and $E + dE$. The function $g(E)$, called the *density of states*, is the *number of states available per unit energy range*. The function $n(E)/g(E)$ is *distribution function*. A comparison of Eqs. (2.63) and (2.64) and replacing e^α by A gives

$$F_{FD}(E) = \frac{1}{A \exp\left(\frac{E}{kT}\right) + 1} \quad \dots(2.65)$$

The function $F_{FD}(E)$ is known as *Fermi-Dirac distribution function*. The constant A is strongly dependent on temperature and the dependence is exponential, given by

$$A = \exp(-E_F/kT) \quad \dots(2.66)$$

Therefore, the Fermi-Dirac distribution is given by

$$F_{FD}(E) = \frac{1}{\exp[(E - E_F/kT) + 1]} \quad \dots(2.67)$$

where, E_F is called Fermi energy and is a constant for a given system. The variation of $F_{FD}(E)$ with energy E at three different temperatures is shown in Fig. 2.3. The figure shows that at $T = 0K$, all quantum states with energy below E_F are completely filled and those above E_F are completely empty. Consider the following cases.

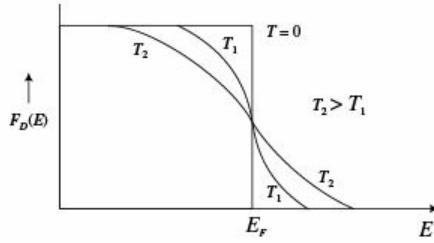


Fig. 2.3 Fermi-Dirac distribution versus energy at three different temperatures.

- (a) At $T = 0\text{ K}$: When $E < E_F$ so that $E - E_F$ is negative, then

$$F_{FD}(E) = \frac{1}{e^{-\infty} + 1} = \frac{1}{0 + 1} = 1$$

$\therefore F_{FD}(E) = 1$, which implies all quantum states are completely occupied.

When $E > E_F$ so that $E - E_F$ is positive, then

$$F_{FD}(E) = \frac{1}{e^{+\infty} + 1} = \frac{1}{\infty} = 0$$

Which implies that all quantum states with $E > E_F$ are empty. Thus, we can define Fermi energy as energy of the uppermost occupied level at 0 K .

- (b) At $T > 0\text{ K}$: At $T > 0\text{ K}$ electrons are excited to vacant levels above E_F (generally $E_F \gg kT$), resulting in few quantum states below E_F (within energy range kT) becoming empty and few quantum states above E_F (within energy range kT) being occupied. Thus, the probability of finding a particle (e.g., electrons in a metal) in a state with $E > E_F$ becomes greater than zero and the probability for the states with $E < E_F$ becomes less than unity.

$$\text{At } T > 0\text{ K and } E = E_F: F_{FD}(E) = \frac{1}{e^0 + 1} = \frac{1}{1 + 1} = 0.5$$

Hence, the probability of occupancy of E_F at $T > 0$ is always 0.5. Thus, we can also define

Fermi energy as the energy of the state at which the probability of occupation is $\frac{1}{2}$ at any temperature above 0 K . At $E \gg E_F$

$$F_{FD}(E) = \frac{1}{\exp[(E - E_F)/kT] + 1} \approx \frac{1}{\exp(E - E_F)/kT}$$

which is Maxwell-Boltzmann distribution.

2.13 ENERGY OF A FERMI-DIRAC SYSTEM

Electrons in a metal are loosely bound and they obey Fermi-Dirac statistics. Therefore, the electrons in a metal can be treated as a gas of free particles. The Fermi-Dirac distribution law, given by Eq. (2.63), may be expressed in the form

$$dN = \frac{g(E)dE}{A \exp\left(\frac{E}{kT}\right) + 1} \quad \dots(2.68)$$

where, dN is the number of electrons in the energy range E and $E + dE$, $g(E)$ is the number of eigenstates in the energy range E and $E + dE$ and $A = e^\alpha$. The value of $g(E) dE$, for the two different spin states is given by

$$g(E)dE = \frac{8\pi m V}{h^3} (2mE)^{1/2} dE \quad \dots(2.69)$$

Therefore, Eq. (2.68) can be written as

$$dN = \frac{8\pi mV}{h^3} \frac{(2m)^{1/2} dE}{A \exp\left(\frac{E}{kT}\right) + 1}$$

Integrating this equation, we have

$$N = \frac{8\pi mV}{h^3} (2m)^{1/2} \int_0^\infty \frac{E^{1/2} dE}{A \exp(E/kT) + 1} \quad \dots(2.70)$$

Total energy is given by

$$E(T) = \int E dN = \frac{8\pi mV(2m)^{1/2}}{h^3} \int_0^\infty \frac{E^{3/2} dE}{A \exp\left(\frac{E}{kT}\right) + 1} \quad \dots(2.71)$$

Evaluating this integral is complicated. However, at $T=0$ K, the Fermi-Dirac distribution function (Eq. (2.65)) is equal to unity for $E < E_F$ and zero for $E > E_F$. At $T=0$ K, the energy of the highest occupied level is called the Fermi energy. Considering, Fermi temperature T_F given by

$$T_F = \frac{E_F}{k} \quad \dots(2.72)$$

At $T=0$ K, Eqs. (2.70) and (2.71) reduce to

$$N = \frac{8\pi mV(2m)^{1/2}}{h^3} \int_0^{E_F} E^{1/2} dE \quad \dots(2.73)$$

$$E(0) = \frac{8\pi mV(2m)^{1/2}}{h^3} \int_0^{E_F} E^{3/2} dE \quad \dots(2.74)$$

where $E(0)$ is the total energy at 0 K. Evaluating the integral in Eq. (2.73) and solving for E_F , we get

$$E_F = \frac{\hbar^2}{2m} \left(\frac{3N}{8\pi V} \right)^{2/3} = \frac{\hbar^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3} \quad \dots(2.75)$$

The quantity N/V can be obtained from Hall effect measurement and so E_F can be calculated. The Fermi speed v_F is defined such that

$$E_F = \frac{1}{2} m v_F^2 \quad \dots(2.76)$$

Evaluation of integral in Eq. (2.74) and replacing E_F by the expression in Eq. (2.75), we get

$$E(0) = \frac{3Nh^2}{10m} \left(\frac{3N}{8\pi V} \right)^{2/3} \quad \dots(2.77)$$

It can easily be seen from Eqs. (2.77) and (2.75) that

$$E(0) = \frac{3}{5} N E_F$$

The mean energy of an electron at absolute zero $E_m(0)$ is given by

$$E_m(0) = \frac{3}{5} E_F \quad \dots(2.78)$$

This energy is called *zero point energy*.

Example 2.1

The density of silver is 10.5 g/cm^3 and its atomic weight is 108. If each atom contributes one electron for conduction, calculate (i) the number density of conduction electrons, (ii) Fermi energy, and (iii) mean energy of the electron at 0 K.

Solution.

(i) Number density of conduction electrons

$$\begin{aligned} \frac{N}{V} &= \frac{\rho N_A}{M} = \frac{10.5 \text{ g/cm}^3 \times 6.02 \times 10^{23} \text{ atoms/mole}}{108 \text{ g/mole}} \\ &= 5.85 \times 10^{22} \text{ cm}^{-3} = 5.85 \times 10^{28} \text{ m}^{-3} \end{aligned}$$

$$\text{(ii) Fermi energy} \quad E_F = \frac{\hbar^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3}$$

$$\begin{aligned} &= \frac{(6.626 \times 10^{-34} \text{ Js})^2}{8 \times 9.1 \times 10^{-31} \text{ kg}} \left(\frac{3}{\pi} \times 5.85 \times 10^{28} \text{ m}^{-3} \right)^{2/3} \\ &= 8.816 \times 10^{-19} \text{ J} = 5.51 \text{ eV} \end{aligned}$$

$$\text{(iii) Mean energy of the electron at } 0 \text{ K} = \frac{3}{5} E_F$$

$$\begin{aligned} &= \frac{3}{5} \times 5.51 \text{ eV} = 3.31 \text{ eV.} \end{aligned}$$

Example 2.2

The density of the sodium is 0.97 g cm^{-3} and its atomic weight is 23. If it has one free electron per atom, calculate its Fermi energy.

Solution.

$$\text{Electron density} \quad \frac{N}{V} = \frac{\rho N_A}{M} = \frac{0.97 \text{ g/cm}^3 \times 6.02 \times 10^{23} \text{ atoms/mole}}{23 \text{ g/mole}}$$

$$\begin{aligned}
 &= 2.53 \times 10^{22} \text{ electron cm}^{-3} \\
 &= 2.53 \times 10^{28} \text{ electrons m}^{-3} \\
 \text{Fermi energy} \quad E_F &= \frac{\hbar^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3} \\
 &= \frac{(6.62 \times 10^{-34} \text{ Js})^2}{8 \times 9.1 \times 10^{-31} \text{ kg}} \times \left(\frac{3}{\pi} \times 2.53 \times 10^{28} \text{ m}^{-3} \right)^{2/3} \\
 &= 5.032 \times 10^{-19} \text{ J} \\
 &= \frac{5.032 \times 10^{-19}}{1.6 \times 10^{-19}} \text{ eV} = 3.145 \text{ eV}
 \end{aligned}$$

SHORT ANSWER TYPE QUESTIONS

1. Differentiate between classical and quantum statistical mechanics.
2. Differentiate between microscopic and macroscopic systems.
3. Name any three macroscopic variables.
4. Define: (a) phase space and (b) ensemble.
5. State and explain Liouville's theorem.
6. When an ensemble is said to be in equilibrium?
7. State and explain postulate of equal probability.
8. State and explain Maxwell-Boltzmann distribution law.
9. Use Maxwell speed law to explain (a) why there is almost no hydrogen in the earth's atmosphere and (b) smaller astronomical bodies have no atmosphere.
10. State the law of equipartition of energy.
11. Differentiate between symmetric and antisymmetric wave functions. Also mention the statistics the particles of which are described by them.
12. Write Bose-Einstein distribution law. Explain Bose-Einstein condensation.
13. Write Fermi-Dirac distribution law and mention its important features.
14. Write the expression for Fermi-Dirac distribution function and show its variation with energy at different temperatures.
15. Define Fermi energy.

LONG ANSWER TYPE QUESTIONS

1. Obtain Maxwell-Boltzmann distribution law.
2. Derive Maxwell's speed distribution law and speed distribution function. Use this to obtain the expressions for most probable speed, mean speed and root mean square speed.
3. Obtain Bose-Einstein distribution law and discuss its important features.
4. Derive Fermi-Dirac distribution law and discuss its important features.

5. Show that the mean energy of an electron at absolute zero is $\frac{3}{5}E_F$.

NUMERICALS

1. Calculate the Fermi energy of copper if the number density of conduction electron is $5.9 \times 10^{28} \text{ m}^{-3}$.
[Ans. 7.03 eV]
2. The density and atomic weight of silver is 10.5 g cm^{-3} and its molecular weight is 107.9 g/mol . Calculate its (a) Fermi energy and (ii) Fermi temperature.
3. The number density of conduction electrons in aluminium is $1.18 \times 10^{28} \text{ m}^{-3}$. Calculate its (a) Fermi energy and (b) mean energy of the electron at 0 K.
[Ans. (a) 11.69 eV, (b) 7.01 eV]
4. The number density of conduction electrons in gold is $5.9 \times 10^{28} \text{ cm}^{-3}$. Calculate its (a) Fermi energy and (b) mean energy of the electron at 0 K.
[Ans. (a) 5.54 eV, (b) 3.32 eV]
5. The Fermi energy of silver is 5.51 eV. What is the average energy of a free electron at 0 K in silver?
[Ans. 3.306 eV]

Unit III

3

Nuclear Physics

3.1 INTRODUCTION

It was famous Rutherford's alpha-particles scattering experiment which proved the existence of *nucleus*. As some of the alpha-particles got scattered through angles greater than 90° and a few of them even rebounded (i.e., got scattered through an angle of 180°), this lead Rutherford and coworkers to believe that *whole of the positive charge and nearly the entire mass of the atom was concentrated in an extremely small central core (of the size $\approx 10^{-15}$ m) called nucleus*. The nucleus is surrounded by suitable number of electrons such that the total negative charge of electrons is equal to the positive charge on the nucleus and thus the atom on the whole is electrically neutral.

3.2 STRUCTURE OF NUCLEUS

Rutherford's experiment proved that nucleus has a very compact structure and practically the whole mass and the total positive charge of the atom is concentrated in the nucleus. Natural radioactivity revealed that the nuclear structure must be very complex, capable of ejecting different particles like α -particles and electrons as well as electromagnetic radiation in the form of γ -rays. The studies of the α -rays and γ -rays indicated that there exists a great order in the internal structure of the nucleus. Despite complexity, the interpretation of the continuous nature of the β -rays spectra introduced a new particle, called *neutrino*, which though having practically no mass plays the important role of discrete carrier of energy.

After the existence of the nucleus was established, several theories were proposed regarding the constituents and structure of the nucleus, e.g., *proton-electron theory*, *proton-neutron theory*, *neutron-positron theory*, etc. In the proton-electron model, it was postulated that the nucleus contained A protons and $(A-Z)$ electrons (A is the mass number and Z is the atomic number). Under this assumption, the nuclear mass becomes A times the proton mass and the charge of the nucleus becomes $A (+e) + (A-Z) (-e) = Ze$, in agreement with experiment. However, this model led to the number of difficulties and finally it was discarded. With the discovery of neutrons in 1932 by James Chadwick, the dilemma came to an end and it was accepted that the nucleus consists of Z protons

and $A-Z$ neutrons, i.e., *the proton-neutron theory got the general acceptance*. Although, the protons and neutrons are themselves made up of quarks but we will treat proton and neutron as elementary particles.

Proton-Neutron Theory: According to this theory, nuclei are composed of protons and neutrons. For example, helium (${}_2^4\text{He}$ with general representation ${}_Z^A\text{X}$) contains two protons and two neutrons. Similarly, ${}_{12}^{24}\text{C}$ denotes a carbon atom with 6 protons and 6 neutrons, while ${}_{92}^{238}\text{U}$ denotes a uranium atom with 92 protons and 146 (= 238 – 92) neutrons.

According to this theory, the total number of protons and neutrons is given by A , the number of protons by Z and the number of neutrons by $A-Z$. The number of electrons outside the nucleus, in a normal neutral atom is evidently Z . However, nuclei are observed to emit electron and positron (β -decay). It was explained that, *the electron does not pre-exist in the nucleus but is formed just at the instant of emission caused by the transformation of a neutron into a proton ($n \rightarrow p + e^-$)*. Similarly, the *positron emission* is due to the converse process, i.e., when a proton transforms itself into a neutron ($p \rightarrow n + e^+$).

As protons and neutrons can be converted into each other in the nucleus, they are regarded as two alternative states of a single heavy nuclear particle to which the name *nucleon* has been given. Depending upon the relative numbers of protons and neutrons, atomic nuclei are classified into following categories.

- (i) **Isotopes:** Nuclei having different mass number A but the same atomic number Z are called isotopes. For example, hydrogen has three isotopes ${}_1^1\text{H}$, ${}_1^2\text{H}$, and ${}_1^3\text{H}$; while isotopes of oxygen are ${}_{16}^{16}\text{O}$, ${}_{17}^{17}\text{O}$ and ${}_{18}^{18}\text{O}$. Naturally occurring elements are often mixture of several isotopes, e.g., of the several isotopes of carbon (${}_{12}^{12}\text{C}$, ${}_{13}^{13}\text{C}$, ${}_{14}^{14}\text{C}$, ${}_{15}^{15}\text{C}$, ${}_{16}^{16}\text{C}$), about 98.9 per cent is ${}_{12}^{12}\text{C}$ while the other isotopes together constitute 1.1 per cent. Since the essential individuality of an atom is determined by the nucleus, the number of naturally occurring elements is much larger than the elements of the periodic table. The factor that distinguishes the different isotopes is the relative number of neutrons ($A-Z$) which vary from isotope to isotope.
- (ii) **Isobars:** Atomic nuclei having the same mass number A but different atomic number Z are called isobars. For example, ${}_{11}^{23}\text{Na}$ and ${}_{10}^{23}\text{Ne}$; ${}_{20}^{40}\text{Ca}$ and ${}_{18}^{40}\text{Ar}$. Also, ${}_{16}^{14}\text{C}$, ${}_{17}^{14}\text{N}$ and ${}_{18}^{14}\text{O}$ are isobars with $A = 14$. In isobars, both the number of protons and number of neutrons are different.
- (iii) **Isotones:** Nuclei having the same number of neutrons are called *isotones*. For example, ${}_{16}^{13}\text{C}$ and ${}_{17}^{14}\text{N}$ are *isotones* as the number of neutrons in their nuclei is seven.
- (iv) **Isomers:** Certain nuclei though having the same A , Z and N (number of neutrons) are distinguishable by certain differences in the internal structure of the nuclei due to different internal energies, which may be manifested by different decay periods in case of radioactive nuclei. Such nuclei are called *isomers*. For example, a nucleus in the excited state (denoted by X^* if the normal nucleus is X) with long lifetime is an isomer.

3.3 GENERAL PROPERTIES OF NUCLEUS

Some of the important properties of the atomic nucleus are discussed below.

- (i) **Nuclear Mass:** The masses of nuclei are expressed in *atomic mass unit (u)*. In this mass scale, a neutral ${}_{12}^{12}\text{C}$ atom is taken as 12.0 u . Thus, *one atomic mass unit (u) is one-twelfth the mass of ${}_{12}^{12}\text{C}$ atom*. In other words

$$1u = \frac{\text{mass of } {}^1_6\text{C atom}}{12} = 1.66043 \times 10^{-27} \text{ kg}$$

The mass of a neutral hydrogen atom is $1.007825 u$ and that of a proton and neutron is $1.007276 u$ and $1.008665 u$ respectively.

The mass of a nucleus is equal to the sum of the masses of neutrons and protons in the nucleus. The mass of a ${}^{12}_6\text{C}$ nucleus is $12 u$ and that of uranium ${}^{238}_{92}\text{U}$ is $238 u$.

- (ii) **Nuclear Charge:** A nucleus is positively charged due to protons contained in it. The number of unit of positive charge carried by a nucleus is equal to the number of its protons. Thus, the charge carried by a nucleus is Ze , where Z is its atomic number and $e = +1.6 \times 10^{-19} \text{ C}$ for a proton. For example, charge on a hydrogen nucleus ($Z = 1$) is $+e$ and that on a ${}^{12}_6\text{C}$ nucleus ($Z = 6$) is $+6e$.
- (iii) **Nuclear Size:** Rutherford experiments established that nuclei are much smaller than atoms, with radii just a few femtometers or fermis ($1 \text{ fm} = 10^{-15} \text{ m}$). It was believed that nuclei are spherical in shape. However, more recent experiments indicate that though majority of the nuclei are spherical, some of them are non-spherical, ranging from prolate (elongated like an American football) to oblate (flattened like pumpkin). But most often the difference between major and minor radii is small and for all practical purposes the nuclei can be treated as spherical.

The radii of nuclei increase steadily from about 2 fm for helium to about 7 fm for uranium. It is found that radii of nuclei approximately follow the relation

$$R = R_0 A^{1/3} \quad (3.1)$$

Here, A is the mass number of a nucleus and R_0 is constant whose value is $R_0 = 1.07 \text{ fm}$. Thus, knowing the mass number of a nucleus, its radius can be estimated using the above relation. For example, for ${}^{12}_6\text{C}$, $R = 1.07 \times 10^{-15} \times (12)^{1/3} = 3.2 \times 10^{-15} \text{ m}$.

We can note from Eq. (3.1) that the volume of the nucleus is proportional to $V \propto R^3 \propto A$. This means that the *volume per nucleon is approximately the same in all nuclei. In other words, the density of nucleons inside the nucleus is the same in all nuclei of the order of $3 \times 10^{17} \text{ kg/m}^3$* , about 10^{14} times greater than that of most liquids and solids. The nucleons are closely packed inside a nucleus, like the molecules in a liquid. This resemblance led to the liquid drop model of nucleus, to be discussed later in this chapter.

- (iv) **Nuclear Quantum States:** Like atoms, nuclei are found to have quantized energy levels and the corresponding set of quantum states. However, the study of α -and γ -ray spectra as well as artificial radioactivity shows that the spacing of the levels (energy difference) in nuclei is many times larger than that in atoms. In heavy nuclei, the level spacing is usually several tens of keV, while in light nuclei it can be up to 10 MeV or more.
- If a nucleus is raised to one of its excited states, (e.g., when hit by energetic neutron) it returns to the ground state by emitting a photon. This is similar to the emission of photons in atomic transitions. However, the energy of the photons is much larger than that emitted in atomic transitions. Photons emitted in nuclear transitions are called γ -rays.

There is a fundamental difference in the nature of force acting in the nuclear region and extra-nuclear region in an atom. In the extra-nuclear region, electrostatic Coulomb's force dominates while inside the nucleus the short range nuclear force reign supreme.

- (v) **Angular Momentum of the Nucleus:** Like the electron, both the neutron and proton have half spin. The total angular momentum \mathbf{J} of a nucleus is the sum of the orbital angular momenta of all its nucleons plus all of their spins. In other words, the total angular momentum of a nucleus arises from the intrinsic spin angular momenta of its protons and neutrons and the total angular momenta due to the motion of these nucleons within the nucleus.

The magnitude of total angular momentum \mathbf{J} is quantized given by

$$J = |\mathbf{J}| = \sqrt{j(j+1)} \hbar \quad (3.2)$$

where j is the total nuclear angular momentum quantum number, usually called the *nuclear spin*. It is important to note that in nuclear physics, the word *spin* usually refers to the total *angular momentum of a nucleus*. The quantum number j is characteristic of a nucleus and can have half integral or integral values including zero. The spin j can be correlated with the Z and A values of the nucleus:

- (i) If the mass number A is odd and Z is even or odd, j is half integer. Examples: ${}^1_1\text{H}$, ${}^{13}_6\text{C}$, ${}^{19}_9\text{F}$, ...
- (ii) If A is even and Z is odd, j has integral value. Examples: ${}^1_2\text{H}$, ${}^{10}_5\text{B}$, ${}^{14}_7\text{N}$, ...
- (iii) For nuclei with both A and Z even, j is zero. Examples ${}^{12}_6\text{C}$, ${}^{16}_8\text{O}$, ${}^{32}_{16}\text{S}$, ...
- (vi) **Magnetic Moment:** The total (or spin) angular momentum of a nucleus gives rise to a magnetic moment to it (though very small compared to that associated with electrons), is given by

$$\mu = g_N \mu_N \mathbf{J}$$

where g_N is the *nuclear g-factor*. It is a dimensionless number whose value is different for different nuclei.

The term μ_N represents *nuclear magneton* defined as

$$\mu_N = \frac{e\hbar}{2m_p} = 5.051 \times 10^{-27} \text{ JT}^{-1} \quad (3.3)$$

where m_p is the mass of a proton. The nuclear magnetic moments are measured in nuclear magnetons. The experimentally measured values of magnetic moment of proton (μ_p) and neutron (μ_n) are

$$\mu_p = 2.79285 \mu_N \text{ and } \mu_n = -1.913041 \mu_N$$

The positive sign of the moment means that the spin angular momentum and the magnetic moment are in the same direction while the negative sign indicate that they are in opposite direction. From the definition of μ_N , one may expect the magnetic moment of proton to be $1\mu_N$ instead of $2.79285\mu_N$. The difference between the expected and the actual value indicates that the proton has a non-uniform charge distribution. The negative magnetic moment of neutron indicates that the charge distribution is complex, containing equal quantities of negative and positive charges.

The magnetic moment of atom is measured in units of Bohr magneton $\mu_B \left(= \frac{e\hbar}{2m_e} \right)$. The ratio of Bohr magneton to nuclear magneton is

$$\frac{\mu_B}{\mu_N} = \frac{m_p}{m_e} = 1/1836 \quad (3.4)$$

Thus, the nuclear magneton is 1836 times smaller than the Bohr magneton. Therefore, if there were electrons in the nucleus, their large magnetic moment would dominate and the total magnetic moment would be of the order of μ_B . However, the measured magnetic moments of all nuclei are much smaller than this and are of the order of μ_N ; confirming the validity of proton-neutron theory.

For deuteron, consisting of a proton and a neutron, the nuclear spin is 1, i.e., the spins of two nucleons are parallel. In this state, the magnetic moment of deuteron would be just the sum of the magnetic moments of proton and neutron given by $(+2.7928 \mu_N - 1.9130 \mu_N) = 0.8798 \mu_N$. The experimentally measured value of μ is $0.857 \mu_N$, lower by nearly 3% from the expected value. The difference is believed to be due to non-spherical nuclear charge distribution and mixing of other states.

3.4 NUCLEAR FORCE

The force that holds electrons in their atomic orbits is the electrostatic force of attraction between the negatively charged electrons and the positively charged nucleus. However, this force cannot hold the nucleons together in the nucleus. First, because the neutrons are uncharged particles and so cannot feel electrostatic force and, second, the protons being positively charged would repel each other. The gravitational forces though acting on all particles is too weak (10^{-36} times weaker than electrostatic force) to hold the nucleons together.

Since neither the electrostatic force nor the gravitational force holds the nucleons together in a nucleus, there must be some other force which holds the nucleons together. The force responsible for holding the nucleons together is called the *nuclear force*. The nuclear force must be stronger than the electrostatic force so that it overcomes the repulsive force between the protons. For this reason, the nuclear force is called the *strong force*.

It is estimated that the minimum kinetic energy of any one nucleon in a nucleus is of the order of 20 MeV. A nucleus to hold together its total energy must be negative. Therefore, the nuclear force on any nucleon due to all other nucleons must produce a potential energy well of depth at least several tens of MeV. While the nuclear potential well must be very deep, it must also be of very short range, i.e., the force between the nucleons must fall rapidly to zero as the separation increases. This is indicated in the experiments in which a beam of protons is scattered by nuclei. Those protons that pass by a nucleus at a large distance are influenced by only Coulomb repulsion between the proton's charge and the nuclear charge. Even the protons passing at a distance of 2 to 3 fm of the nuclear surface, scatter in a manner predictable on the basis of simple Coulomb repulsion. Only for the protons passing at a distance of 1 to 2 fm from the nuclear surface, the scattering patterns change indicating the influence of a strong attractive nuclear force. Thus, it is concluded that the nuclear force between the nucleons is negligibly small when they are more than 2 fm apart.

When the separation between the two nucleons is between 1 and 2 fm, the force between them is strongly attractive. However, as the separation is reduced below 1 fm the nuclear force becomes increasingly repulsive. The constancy of the density inside a nucleus is indicative of the fact that the *nuclear force is not purely attractive*. If it had been so, one would expect the nucleons at the centre of a large nucleus to be forced inward more tightly than that in a small nucleus. This is not what is observed and the density of all nuclei is the same.

Figure 3.1 shows the variation of potential energy of the nuclear force between two nucleons, as a function of separation r between them. The potential energy approaches zero rapidly for separation beyond 2 fm. For separation below 2 fm, the force is attractive and the potential energy becomes negative, dropping to a minimum of about 100 MeV at about 1 fm. As the separation decreases below 1 fm, the attractive force reduces and finally becomes repulsive.

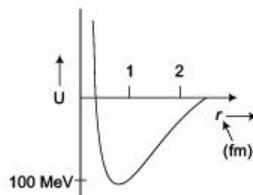


Fig. 3.1 Variation of nuclear potential energy between two nucleons.

Charge Independence of Nuclear Forces: Nuclear force between two nucleons is the same whether they are both protons, or both neutrons, or one proton and one neutron. This property is known as the *charge independence of nuclear forces*, i.e., the nuclear force on a nucleon is independent of the fact that whether the nucleon is charged (a proton) or uncharged (a neutron). It is charge independence that is responsible for the tendency of the nuclei to have equal number of neutrons and protons. The studies of interaction between two nucleons ($p-p$, $p-n$, $n-n$) in collision experiments have revealed the charge independent behaviour of the nuclear force.

Nuclear Force and Electrons: The nuclear force does not act on electrons like electrostatic force does not act on electrically neutral particles. The evidence for this comes from experiments in which high energy electrons collide with nuclei, even when the electrons penetrate deep into the nucleus, there are no evidence that they are affected by the strong nuclear force. This makes the electron a very useful probe of nuclei since its interaction with the nucleons is just the well-understood electromagnetic interaction.

3.4.1 Meson Theory of Nuclear Forces

In 1935, a Japanese scientist, Hideki Yukawa predicted the existence of a particle with mass of the order of $100 \text{ MeV}/c^2$. Because this mass is in between the mass of a nucleon (about $940 \text{ MeV}/c^2$) and an electron ($0.5 \text{ MeV}/c^2$), Yukawa named the particle as *meson*, meaning the particle of intermediate mass. Later several particles of medium mass were discovered and they all are referred to as mesons and the particle that Yukawa referred was named π meson or pion.

Based on the idea that every force in nature is ‘carried’ by a particle, he argued that the strong nuclear force must be ‘carried’ by a pion just as electromagnetic force is carried by the photon. In

1947, the pion was discovered with almost the same mass as predicted by Yukawa. Yukawa was awarded Nobel Prize in physics for his prediction.

Yukawa's idea was based on the relation of photon to electromagnetic forces. Energy in electromagnetic fields is carried by photons. In a sense, therefore, it can be said that electromagnetic forces are carried by photons. The exact sense in which photons carry the electromagnetic force emerges from a complete theory of the electromagnetic field. The quantum field theory developed by Dirac in 1927. In a simplified way, it emerges from the quantum field theory that the electromagnetic forces between two charged particles result from one particle emitting a photon and the second particle absorbing this photon. This exchange of a photon between two electrons or any other charged particles, can be pictorially depicted by a Feynmann diagram shown in Fig. 3.2. The emission of a photon by the left-hand electron changes the electron's momentum and the subsequent absorption of the photon by the right-hand electron changes its momentum by an equal and opposite amount. This exchange of momentum leads to the generation of force between the two electrons.

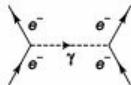


Fig. 3.2 Feynmann diagram for phonon exchange between two electrons.

Yukawa argued that if electromagnetic forces are carried by a particle (photon), the strong nuclear forces may also be carried by exchange of a particle, i.e., the strong force between two nucleons might be the result of the nucleons exchanging a particle. To explain the known properties of the nuclear force between all possible pairs of nucleons ($p-p$, $p-n$, $n-n$), it was found that there must be three kinds of exchange particles. These are now called *pions* and are denoted as π^+ , π^0 , π^- , indicating their charges of $+e$, 0 and $-e$. Figure 3.3 shows the three possibilities. In (a) two protons exchange a π^0 , in (b) a proton and a neutron exchange π^+ and in (c) a neutron and a proton exchange a π^- . In the last two processes the emission and absorption of a charged pion changes a proton into a neutron and vice versa.

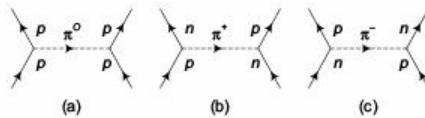


Fig. 3.3 Feynmann diagram for pion exchange between two nucleons.

Yukawa's prediction seems to be satisfactory. At the time of Yukawa prediction no particle with rest mass between an electron ($0.5 \text{ MeV}/c^2$) and that of a proton ($\approx 938 \text{ MeV}/c^2$) was known. However, later π -meson with rest masses

$$m_{\pi^+} = m_{\pi^-} = 139.6 \text{ MeV}/c^2$$

and

$$m_{\pi^0} = 135.0 \text{ MeV}/c^2$$

were discovered. Also, from the symmetry of neutron-proton scattering experiments, it was concluded in approximately half the scatterings, the neutron changes into a proton and the proton changes into a neutron while in the other half, the neutron and proton do not exchange identities

when they interact. The two ways in which the first half exchange happen can be represented as

$$n \rightarrow p + \pi^- \text{ followed by } \pi^- + p \rightarrow n$$

and

$$p \rightarrow n + \pi^+ \text{ followed by } \pi^+ + n \rightarrow p$$

In the above cases, a meson transfers charge as well as momentum between the two nucleons. In other half of scatterings, the exchanges can be represented as

$$n \rightarrow n + \pi^0 \text{ followed by } \pi^0 + p \rightarrow p$$

and

$$p \rightarrow p + \pi^0 \text{ followed by } \pi^0 + n \rightarrow n$$

In these cases, the neutron and proton do not exchange charge when they interact. The exchanged meson carries the transferred momentum.

3.5 NUCLEAR BINDING ENERGY–STABILITY OF NUCLEI

The nucleons of any nucleus are bound together by strong nuclear forces making the nucleus a stable structure. A definite amount of energy has to be supplied to break a nucleus into its constituents. This energy is referred to as the binding energy of the nucleus. *The binding energy of a nucleus is defined as the energy required to break the nucleus into its constituent nucleons apart such that they do not interact with each other.* In other words, the binding energy of a nucleus is the energy with which nucleons are bound in the nucleus. The binding energy of a nucleus is a measure of its stability, higher is the binding energy more stable is the nucleus.

There is always a difference between the mass of a nucleus and the sum of the masses of its nucleons (Z protons and N neutrons), often called the ‘mass defect’. This mass defect appears as the binding energy of the nucleus in accordance with Einstein’s mass energy relation ($E = \Delta m \cdot c^2$).

For a nucleus ${}_Z^A X$, containing Z protons and $N (= A - Z)$ neutrons, the binding energy E_b is given by

$$E_b = [Zm_p + Nm_n - m({}_Z^A X)] \cdot c^2 \quad (3.5)$$

where m_p is the mass of a proton, m_n the mass of a neutron, $m({}_Z^A X)$ is the mass of the nucleus.

Most of the times, the tables of masses usually list the atomic mass m_{atom} rather than the nuclear mass, since it is m_{atom} that is usually measured. The mass m_{atom} is equal to m_{nuc} plus the mass of Z atomic electrons (provided we neglect the binding energy of the electrons, although this do not cause a large error because electron binding energies are in k eV compared to nuclear binding energies in MeV). Thus,

$$m_{\text{atom}} = m_{\text{nuc}} + Z m_e$$

Now, we can add $Z m_e$ to $m_{\text{nuc}} ({}^A_Z X)$, if we also add $Z m_e$ to the term $Z m_p$. Replacing $Z(m_p + m_e)$ by $Z m_H$, where $m_H (= m_p + m_e)$ is mass of a hydrogen atom. The expression for the nuclear binding energy can be written as

$$E_b = [Z m_H + N m_n - m_{\text{atom}}] \cdot c^2 \quad (3.6)$$

Using a mass spectrometer, the atomic masses can be calculated accurately which can then be used to calculate the nuclear binding energies using Eq. (3.6). The direct measurement of binding energies, by complete separation of a nucleus, is seldom possible.

Very often, the average binding energy per nucleon $\left(\frac{E_b}{A}\right)$ is used in the discussion rather than

the total binding energy. Figure 3.4 shows the variation of average binding energy per nucleon as a function of the atomic mass number A . The characteristic features of the curve are:

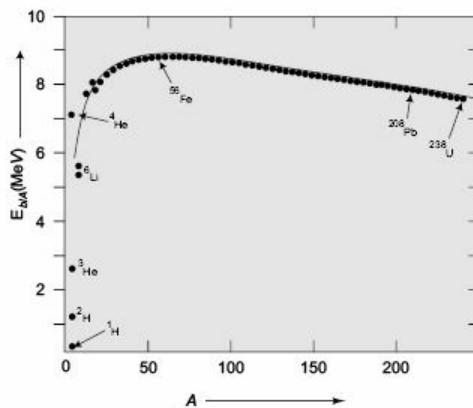


Fig. 3.4 Variation of binding energy per nucleon with mass number.

- For low atomic mass number ($A < 20$), the binding energy per nucleon is very low. As A falls below 20, (E_b/A) falls rapidly to zero for hydrogen, which has no nuclear binding energy. This decrease in E_b/A occurs because almost all nucleons in a small nucleus are close to the surface and the (negative) surface energy is large.
- For most nuclei, the binding energy per nucleon is close to 8 MeV with maximum E_b/A of 8.8 MeV for iron (^{56}Fe).
- Beyond iron, E_b/A decreases slowly with A becoming 7.6 MeV for ^{238}U . This decrease in E_b/A with increasing A is attributed to the increased Coulomb repulsion between the protons of the nuclei.

The lower values of E_b/A when A is large (compared to the nuclei with A near $A = 56$), imply that energy is released when a heavy nucleus splits (or undergoes fission) into two lighter nuclei. The lower values of E_b/A for nuclei with very low A imply that energy is released when they combine (or undergo fusion) to form a heavier nucleus.

Stable Nuclei: A nucleus is stable if it does not transform itself into another nucleus on its own i.e., without the supply of energy from outside. The binding energy of a nucleus is a measure of its stability, larger the binding energy higher is its stability.

Figure 3.5 shows the plot of Z (proton number) versus N (number of neutrons) of stable nuclei (called *line of stability*) along with the $N = Z$ line. It may be noted in the figure, that for light nuclei ($Z \leq 20$), the number of protons and number of neutrons are equal. For nuclei with $Z > 20$, the slope of the line is not 45° , as for $N = Z$ line, i.e., the line deviates from $N = Z$ line towards y -axis. This indicates for nuclei with $Z > 20$, there is an excess of neutrons. For example, the stable isotope of gold ^{197}Au has 79 protons and (197–79 =) 118 neutrons. Thus, the neutron excess is 39. The

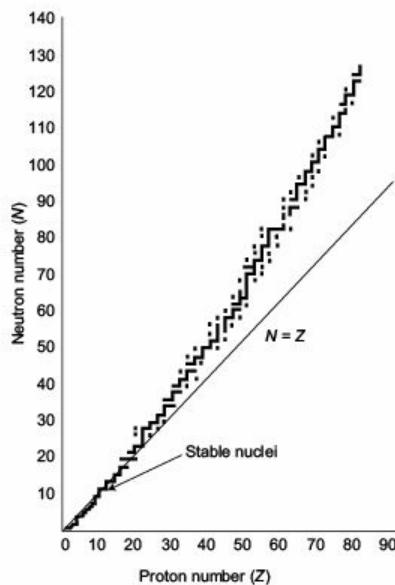


Fig. 3.5 Neutron-proton diagram for stable nuclei.

tendency of the stable nuclei to have more number of neutrons than the protons is as follows. A neutron in a nucleus interacts with other nucleons only through the attractive nuclear force. While a proton, being positively charged, also experiences coulombic repulsive due to other protons in addition to the attractive nuclear force with other nucleons. The repulsive forces increase the energy of the nucleus. This may prompt you to think that the stable nuclei should consist only of neutrons. In fact, the stability of a nucleus (or its binding energy) depends on the proton-neutron ratio and the exact number of protons and neutrons in the nucleus. It has been observed that a nucleus with an even number of both neutrons and protons tends to be the most stable. In Table 3.1 the distribution of nuclei with even/odd pairs of protons/neutrons is given.

Table 3.1 Distribution of stable nuclei

A	Z	N	Number of known stable nuclei
Even	Even	Even	166
Odd	Even	Odd	57
Odd	Odd	Even	53
Even	Odd	Odd	8

Most of the stable nuclei have even Z and even N and there are very few nuclei with odd Z and odd N . This is indicative of the fact that the nuclear force must be such that odd number of particles cannot be tightly bound. In other words, a nucleus in which nucleons of the same kind occur in pairs is highly stable and a nucleus containing an unpaired proton or an unpaired neutron is much less stable.

Example 3.1

Calculate the binding energy and average binding energy per nucleon of $^{12}_6C$. Given: mass of a neutron = 1.008665 u, mass of 1H = 1.007825 u, mass of $^{12}_6C$ = 12.0 u.

Solution. $^{12}_6C$ has 6 protons and 6 neutrons.

$$\text{Mass of 6 neutrons} = 6 \times 1.008665 \text{ u} = 6.05199 \text{ u}$$

$$\text{Mass of 6 hydrogen atoms (mass of 6 protons + 6 electrons)} = 6 \times 1.007825 \text{ u} = 6.04695 \text{ u}$$

$$\text{Total mass of constituents of } ^{12}_6C = 12.09894 \text{ u}$$

$$\therefore \text{Mass defect} = 12.09894 - 12.0 = 0.09894 \text{ u}$$

$$\text{Binding Energy} = 0.09894 \text{ u} \times 931.5 \text{ MeV/u} = 92.16 \text{ MeV}$$

$$\therefore \text{Average binding energy per nucleon} = 92.16/12 \text{ MeV} = 7.68 \text{ MeV.}$$

Example 3.2

Calculate the binding energy and average binding energy per nucleon of the nucleus $^{35}_{17}Cl$. Given: mass of $^{35}_{17}Cl$ atom = 34.969 u, mass of a neutron m_n = 1.0087 u and mass of hydrogen 1H atom m_H = 1.0078 u.

Solution. The $^{35}_{17}Cl$ atom has 17 protons and 18 (= 35 – 17) neutrons. Using the expression for binding energy, we have

$$\begin{aligned} E_b &= [17m_H + 18m_n - m(^{35}_{17}Cl)] \cdot c^2 \\ &= [17 \times (1.0078 \text{ u}) + 18(1.0087 \text{ u}) - 34.969 \text{ u}] \cdot c^2 \\ &= [35.289 \text{ u} - 34.969 \text{ u}] \cdot c^2 \\ &= 0.3294 \times \frac{931 \text{ MeV}}{\text{u}} = 289 \text{ MeV} \\ \frac{E_b}{A} &= \frac{289}{35} = 8.257 \text{ MeV} \end{aligned}$$

Example 3.3

A neutron splits into a proton and an electron. What is the energy produced in this reaction in MeV? Given: mass of an electron = 9.1×10^{-31} kg, mass of a proton = 1.6725×10^{-27} kg, mass of a neutron = 1.6747×10^{-27} kg.

Solution. Mass defect Δm = Mass of neutron – (mass of proton + mass of electron)

$$\begin{aligned} &= 1.6747 \times 10^{-27} \text{ kg} - [1.6725 \times 10^{-27} + 9.1 \times 10^{-31} \text{ kg}] \\ &= 1.6747 \times 10^{-27} \text{ kg} - [(1.6725 + 0.00091) \times 10^{-27} \text{ kg}] \\ &= 1.6747 \times 10^{-27} \text{ kg} - 1.67341 \times 10^{-27} \text{ kg} \\ &= (1.67470 - 1.67341) \times 10^{-27} \text{ kg} = 0.00129 \times 10^{-27} \text{ kg} \end{aligned}$$

$$\begin{aligned}
 \therefore \text{Energy released} &= \Delta m \cdot c^2 \\
 &= 0.00129 \times 10^{-27} \text{ kg} \times (3 \times 10^8 \text{ m/s})^2 \\
 &= 0.00129 \times 10^{-27} \times 9 \times 10^{16} \text{ Joule} \\
 &= 0.01161 \times 10^{-11} \text{ Joule} \\
 &= 1.161 \times 10^{-13} \text{ Joule} \\
 &= \frac{1.161 \times 10^{-13}}{1.6 \times 10^{-13}} \text{ MeV} \\
 &= 0.725 \text{ MeV}
 \end{aligned}$$

Example 3.4

A proton and antiproton annihilate into two photons of the same frequency. What is the wavelength of the photons produced? (Mass of a proton = 1.67×10^{-27} kg)

Solution. Mass of a proton = Mass of a antiproton = 1.67×10^{-27} kg

$$\therefore \text{Mass of proton + antiproton} = 2 \times 1.67 \times 10^{-27} \text{ kg}$$

Energy released in annihilation process, in accordance with Einstein's mass-energy relation, is given by

$$\begin{aligned}
 E &= mc^2 = 2 \times 1.67 \times 10^{-27} \times (3 \times 10^8)^2 \text{ Joule} \\
 &= 2 \times 1.67 \times 9 \times 10^{-11} \text{ J} = 30.06 \times 10^{-11} \text{ Joule}
 \end{aligned}$$

Since annihilation of proton and antiproton gives rise to two photons, hence $2h\nu = E$

$$\begin{aligned}
 \text{or } 2h \left(\frac{c}{\lambda} \right) &= E \Rightarrow \lambda = \frac{2hc}{E} \\
 \text{Wavelength of photon } \lambda &= \frac{2 \times 6.64 \times 10^{-34} \times 3 \times 10^8}{30.06 \times 10^{-11}} \\
 &= \frac{39.78 \times 10^{-15}}{30.06} \text{ m} \\
 &= 1.323 \times 10^{-15} \text{ m.}
 \end{aligned}$$

3.6 DECAY OF UNSTABLE NUCLEI

A nucleus which is unstable decays to form a stable nucleus. Such a nucleus is called a *radioactive nucleus*. The nuclei with $Z \geq 83$ are naturally occurring radioactive nuclei.

In 1896, Henri Becquerel, a French physicist discovered that uranium salts emit radiation that could darken a photographic film, even if the film was wrapped in black lightproof paper. The process established the existence of a radiation which could pass through the black paper. The emission of the new radiation was called *radioactivity*. Within next two years Marie and Pierre Curie had identified three more radioactive elements: thorium, polonium and radium. Becquerel along with Curies shared Nobel Prize in physics for their work on radioactivity. Although, it was not known that the radioactivity originates in the atomic nucleus and the existence of the nucleus could be established only after 15 years of the discovery of radioactivity.

It was later found that the radioactive radiations can be divided into three types, which Ernest Rutherford called α -, β - and γ -rays. We will study in detail about these rays later in this chapter.

3.7 THE RADIOACTIVE DECAY LAW

Rutherford and Soddy, based on their experimental observations, formulated the following laws, often called the laws of *radioactive disintegration*. According to these laws:

- (i) Radioactive decay is a spontaneous process independent of external conditions like temperature, pressure, etc. In a sample of radioactive nuclei, which particular nucleus undergoes disintegration cannot be predicted, it is purely a statistical process.
- (ii) Nuclei decay one by one over a period of time.
- (iii) The rate of disintegration (i.e., number of atoms disintegrated per second) at any instant is directly proportional to the number of radioactive atoms present at that instant. This is known as *radioactive decay law*. This means that as the decay progresses and the original number of atoms go on decreasing, consequently the rate of decay also decreases.

Let us consider a radioactive element containing N_0 nuclei at $t = 0$. Let this number reduces to N after time t due to disintegration of nuclei. Suppose, a small number of nuclei dN disintegrate in time interval between t and $t + dt$. Thus, the rate of disintegration during this time interval is $\frac{dN}{dt}$.

According to the above stated decay law,

$$\frac{dN}{dt} \propto N$$

or

$$\frac{dN}{dt} = -\lambda N \quad (3.7)$$

where λ is a constant of proportionality called *decay or disintegration constant*. The negative sign shows that N is decreasing with increase in time.

On integrating Eq. (3.7) and using the boundary condition that at $t = 0$, $N = N_0$, to evaluate the constant of integration, we get

$$\log_e N = -\lambda t + \log_e N_0$$

or

$$\log \frac{N}{N_0} = -\lambda t$$

or

$$\frac{N}{N_0} = e^{-\lambda t}$$

or

$$N = N_0 e^{-\lambda t} \quad (3.8)$$

Equation (3.8) indicates that the radioactive decay is exponential, as shown in Fig. 3.6. Thus, at $t = \infty$, $N = 0$. In fact the exponential decay law is valid for all unstable system as long as the decay is governed by a chance process.

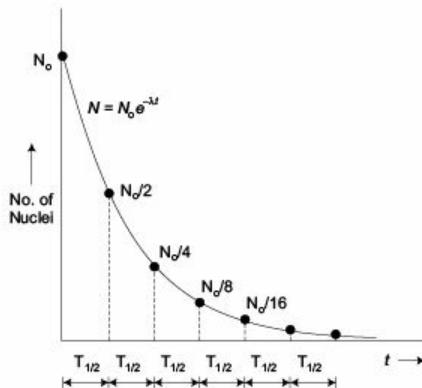


Fig. 3.6 Exponential decay of radioactive nuclei.

Decay Constant: Substituting $t = \frac{1}{\lambda}$ in Eq. (3.8), we get

$$N = N_0 e^{-\lambda t} = N_0 e^{-1} = \frac{N_0}{e}$$

That is, the decay constant can be defined as the reciprocal of the time during which the original number of nuclei (N_0) in a radioactive sample reduces to $\frac{N_0}{e}$, or $\frac{1}{e}$ times the original number of nuclei N_0 in the sample.

Half-life: The exponential decay implies that the time taken for all the radioactive nuclei to decay will be infinite. In other words, it is difficult to find the time a sample of radioactive nuclei will take to decay completely. However, the time taken to decay to half of its initial number can be certainly calculated. This time period is called half-life. The half-life ($T_{1/2}$) of a radioactive element is defined as the time in which half of its nuclei initially present decays.

$$\text{That is, when } t = T_{1/2}, N = \frac{N_0}{2}$$

Substituting this in Eq. (3.8), we get

$$\frac{N_0}{2} = N_0 e^{-\lambda T_{1/2}}$$

or

$$\frac{1}{2} = e^{-\lambda T_{1/2}}$$

or

$$2 = e^{\lambda T_{1/2}}$$

Taking the logarithm of both the sides, we get

$$\lambda T_{1/2} = \log 2 =$$

$$2.3026 \log_{10} 2 = 2.3026 \times 0.3010 = 0.693$$

or $T_{1/2} = \frac{0.693}{\lambda}$ (3.9)

The half-life of radioactive nuclei vary from as short as 10^{-14} sec to as long as 10^{14} years.
Further, after the second half-life

$$N = \frac{(N_0/2)}{2} = \frac{N_0}{4} = N_0 \left(\frac{1}{2}\right)^2$$

After the third half-life, $N = \frac{\left(\frac{N_0}{4}\right)}{2} = \frac{N_0}{8} = N_0 \left(\frac{1}{2}\right)^3$

In general, after the n half-lives,

$$N = N_0 \left(\frac{1}{2}\right)^n = \frac{N_0}{2^n} \quad (3.10)$$

The number of half-lives n can also be put as $n = \frac{t}{T_{1/2}}$,

Therefore,

$$N = \frac{N_0}{2^{\frac{t}{T_{1/2}}}} \quad (3.11)$$

or $N(t) = N_0 2^{-t/T_{1/2}}$ (3.12)

Mean-life: In a sample of radioactive nuclei, a particular nucleus may undergo disintegration in the very beginning or may be the last one to undergo disintegration, i.e., its lifetime is infinite. In other words, the actual life time of radioactive nucleus varies from zero to infinity.

The mean (or average) life of a radioactive nucleus is the average time the nucleus survives before it decays.

As obvious from the definition, the mean life (τ) can be obtained by calculating the total lifetime of all the nuclei initially present (N_0) and dividing it by N_0 .

Let us consider a sample of radioactive element containing N_0 atoms in the beginning, i.e. at $t = 0$. Let the number of nuclei left at time t be N . Assume that dN nuclei disintegrate in small time interval dt . Thus, the lifetime of these dN atoms lies between t and $t + dt$. If dt is infinitesimally small, then lifetime of dN atoms = $t dN$

Thus, the total lifetime of all the nuclei in the sample

$$= \int_0^{N_0} t \cdot dN.$$

By definition, the mean life of a radioactive nucleus

$$\tau = \frac{\int_0^{N_0} t \cdot dN}{N_0} \quad (3.13)$$

Using the relation $\frac{dN}{dt} = -\lambda N$ or $dN = -\lambda N dt$, and

$$N = N_0 e^{-\lambda t}, \text{ we get}$$

$$dN = -\lambda (N_0 e^{-\lambda t}) dt$$

Substituting this value of dN in Eq. (3.13), and changing the limits of integration ($N = 0$ when $t = \infty$ and $N = N_0$ when $t = 0$), we get

$$\tau = \int_{\infty}^0 -\lambda N_0 e^{-\lambda t} dt \cdot t$$

or $\tau = \lambda \int_0^{\infty} t e^{-\lambda t} dt$

or $\tau = \lambda \left[\frac{te^{-\lambda t}}{-\lambda} \Big|_0^{\infty} - \int_0^{\infty} \frac{e^{-\lambda t}}{-\lambda} dt \right]$

or $\tau = \lambda \left[0 + \frac{1}{\lambda} \int_0^{\infty} e^{-\lambda t} dt \right] = \int_0^{\infty} e^{-\lambda t} dt = \left[\frac{e^{-\lambda t}}{-\lambda} \right]_0^{\infty}$

or $\tau = 0 - \frac{1}{-\lambda} = \frac{1}{\lambda}$ (3.14)

Thus, the mean life of a radioactive nucleus is reciprocal of its decay constant.

Using Eq. (3.9), we have

$$\tau = \frac{T_{1/2}}{0.693} = 1.44 T_{1/2} \quad (3.15)$$

That is, the mean life of a radioactive nucleus is 1.44 times its half-lifetime.

Rate of decay: When performing an experiment, it is more convenient to measure the number of radioactive nuclei decaying per second, i.e., *decay rate*, rather than the number of nuclei present at any given time. *The rate of decay is the number of decays (or disintegrations) taking place per unit time.* It is also known as *activity* of the radioactive nuclei and is denoted by R . It is expressed as

$$R(t) = \left| \frac{dN}{dt} \right| = \lambda N(t) \quad (3.16)$$

where $N(t)$ is the number of radioactive nuclei at a time t . As $N(t)$ varies exponentially, $R(t)$ also varies exponentially, or

$$R(t) = R_0 e^{-\lambda t} \quad (3.17)$$

where R_0 is the initial activity at a time $t = 0$ and $R(t)$ that at a time t .

Units of decay rate: The unit used to measure the activity is the *Curie (Ci)*. One Curie is defined as radioactive disintegration rate of 3.70×10^{10} nuclei per second, i.e.,

$$1 \text{ Curie (Ci)} = 3.70 \times 10^{10} \text{ disintegrations/sec.}$$

$$1 \text{ milli Curie (m Ci)} = 3.70 \times 10^7 \text{ decays/sec}$$

$$1 \text{ micro Curie (\mu Ci)} = 3.70 \times 10^4 \text{ decays/sec}$$

This definition has its origin to the fact that one gram of $^{226}_{\text{Ra}}$ undergo 3.7×10^{10} decays per second. Sometimes, one Rutherford ($= 10^6$ decays per second) is also used.

Example 3.5

Calculate the half-life period of a radioactive substance if its activity drops to $\frac{1}{16}$ th of its initial value in 30 years.

Solution. Let the number of half-lives be n , so that

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

or

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

Given

$$\frac{N}{N_0} = \frac{1}{16}$$

∴

$$\frac{1}{16} = \left(\frac{1}{2}\right)^n$$

or

$$\left(\frac{1}{2}\right)^4 = \left(\frac{1}{2}\right)^n \Rightarrow n = 4$$

$$\therefore \text{Half-life } T_{1/2} = \frac{30}{4} = 7.5 \text{ years.}$$

Example 3.6

The nucleus of ^{231}Th is β -radioactive, with half-life $T_{1/2} = 25.6$ hours. At $t = 0$ a sample of ^{231}Th contains $N_0 = 56,000$ nuclei. How many ^{231}Th nuclei will remain after 3 days? What is the total rate at which electrons are emitted at $t = 0$? What will be the rate after 3 days?

Solution. A period of 3 days is about 3 half-lives, and N_0 will get halved three times, i.e.,

$$N = \frac{N_0}{2^3} = \frac{56,000}{8} = 7000$$

A more accurate answer will require the calculation of the number of half-lives in a period of 3 days accurately:

$$n = \frac{t}{T_{1/2}} = \frac{72 \text{ hours}}{25.6} = 2.81$$

$$\therefore N = \frac{N_0}{2^{2.81}} = \frac{56,000}{7.01} \approx 8000$$

The same calculation can also be done in terms of the decay constant λ , as

$$\lambda = \frac{\ln 2}{T_{1/2}} = \frac{0.693}{25.6 \text{ hours}} = 0.0271 \text{ hour}^{-1}$$

With $t = 3$ days = 72 hours, using the exponential decay law

$$N(3 \text{ days}) = N_0 e^{-\lambda t} = 56000 \times e^{-(0.0271 \times 72)} \approx 8000.$$

Each decay of ^{231}Th gives one electron, so the rate of electron emission is the same as the rate R of decays.

We know

$$R(t) = \lambda N(t)$$

At

$$t = 0,$$

$$R_0 = \lambda N_0 = (0.0271/\text{hour}) \times 56,000$$

$$= 1520 \text{ electrons/hour}$$

$$\approx 25 \text{ electrons/minute}$$

As the decay rate $R(t)$ is proportional to the number of nuclei $N(t)$, the $R(t)$ should also have exponential decay as is the case with $N(t)$. Therefore, as N drops to $\frac{1}{7.01}$ to its original value, therefore, the same is true for R :

$$R(3 \text{ days}) = \frac{R_0}{7.01} = \frac{25}{7.01} = 3.6 \text{ electrons/minute.}$$

Example 3.7

The half-life of ^{238}U against α -decay is 4.5×10^9 years. How many disintegrations per second occur in 1 gram of ^{238}U ?

Solution. Half life of

$$^{238}\text{U} = 4.5 \times 10^9 \text{ years}$$

$$= 4.5 \times 10^9 \times 3.16 \times 10^7 \text{ sec} = 1.422 \times 10^{17} \text{ sec.}$$

238 gms of ^{238}U contains 6.023×10^{23} atoms (= Avogadro's number)

$$\therefore 1 \text{ gm of } ^{238}\text{U will contain} = \frac{6.023 \times 10^{23}}{238} \text{ atoms}$$

Using,

$$R = \lambda N$$

$$= \frac{0.693}{T_{1/2}} \times N$$

$$= \frac{0.693}{1.422 \times 10^{17}} \times \frac{6.023 \times 10^{23}}{238} = 1.23 \times 10^4 \text{ decays per sec.}$$

Example 3.8

There is a stream of neutrons with a kinetic energy of 0.0327 eV. If the half-life of neutrons is 700 sec, what fraction of neutrons will decay before they travel a distance of 10 km? Mass of neutron = 1.675×10^{-27} kg.

Solution. Here, kinetic energy $E_k = 0.0327 \text{ eV} = 0.0327 \times 1.6 \times 10^{-19} \text{ J}$

$$\text{As} \quad E_k = \frac{1}{2} mv^2$$

$$\therefore v = \sqrt{\frac{2E_k}{m}} = \sqrt{\frac{2 \times 0.327 \times 1.6 \times 10^{-19}}{1.675 \times 10^{-27}}} = 2.5 \times 10^3 \text{ m/s}$$

Time taken by neutrons to traverse a distance of 10 km

$$= \frac{10 \times 10^3 \text{ m}}{2.5 \times 10^3 \text{ m/s}} = 4 \text{ sec.}$$

The number of half-lives in 4s,

$$n = \frac{t}{T_{1/2}} = \frac{4}{700} = \frac{1}{175}$$

Using, $\frac{N}{N_0} = \left(\frac{1}{2}\right)^n$, we have

$$\frac{N}{N_0} = \left(\frac{1}{2}\right)^{1/175} \approx 0.996$$

$$\therefore \text{Fraction of neutrons decayed} = 1 - \frac{N}{N_0} = 1 - 0.996 = 0.004.$$

Example 3.9

The half-life of ^{210}Po is 138 days. Calculate the amount of polonium (which is α -radioactive) required to produce a source of α -particles of 10 mCi strength.

Solution. Here, decay rate $R = 10 \text{ mCi} = 10 \times 3.7 \times 10^7 \text{ decays/sec.}$

$$= 3.7 \times 10^8 \text{ decays/sec}$$

Half-life $T_{1/2} = 138 \text{ days} = 138 \times 24 \times 60 \times 60 \text{ sec.}$

$$= 1.192 \times 10^7 \text{ sec.}$$

Decay rate $R = \left| \frac{dN}{dt} \right| = \lambda N = \frac{0.693}{T_{1/2}} \times N$

$$\therefore N = \frac{R T_{1/2}}{0.693} = \frac{3.7 \times 10^8 \times 1.192 \times 10^7}{0.693} \text{ atoms}$$

$$= 6.36 \times 10^{15} \text{ atoms}$$

6.023×10^{23} atoms of P_0 weigh 210 grams

$$\therefore 6.36 \times 10^{15} \text{ atoms} \times \frac{210}{6.023 \times 10^{23}} \times 6.36 \times 10^{15} = 2.22 \times 10^{-6} \text{ gm.}$$

Thus, 2.22 μg of polonium is required.

Example 3.10

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 fraction of ^{14}C present with the ordinary carbon isotope ^{12}C . When the organism is dead, its interaction with the atmosphere (which maintains the above equilibrium activity) ceases and its activity begins to drop. From the known half-life (5730 years) of ^{14}C , and the measured activity, the age of the specimen can be approximately estimated. This is the principle of carbon dating used in archaeology. Suppose a specimen from Mohenjodaro gives an activity of 9 decays per minute per gram of carbon. Estimate the approximate age of the Indus-Valley civilization.

Solution. Given, normal activity,

$$R_0 = 15 \text{ decays/minute}$$

Present activity

$$R(t) = 9 \text{ decays/min}$$

Half-life

$$T_{1/2} = 5730 \text{ years}$$

Age(time)

$$t = ?$$

Activity is proportional to the number of radioactive atoms present. Therefore,

$$\frac{N(t)}{N_0} = \frac{R(t)}{R_0} = \frac{9}{15}$$

Also,

$$\frac{N(t)}{N_0} = e^{-\lambda t}$$

$$\therefore e^{-\lambda t} = \frac{9}{15} = \frac{3}{5}$$

$$\text{or } e^{\lambda t} = \frac{5}{3}$$

Taking natural log on both the sides

$$\lambda t = \ln \frac{5}{3} = 2.303 \log_{10} \frac{5}{3} = 2.303 \times 0.2218 = 0.5109$$

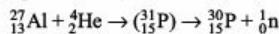
$$\text{or } t = \frac{0.5109}{\lambda} = \frac{0.5109}{0.693} T_{1/2}$$

$$= \frac{0.5109}{0.693} \times 5703 \text{ years} = 4224.3 \text{ years}$$

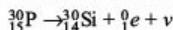
3.8 ARTIFICIAL(OR INDUCED) RADIOACTIVITY

This phenomenon was discovered by Irene Curie Joliot and her husband M. Crederick Joliot. They found that when certain nucleus is bombarded by suitable particles they continue to decay even after bombarding source is removed, i.e., the daughter nucleus becomes radioactive. Thus, *The phenomenon of inducing radioactivity in otherwise stable nuclei by bombarding them with suitable particle is called artificial (or induced) radioactivity.*

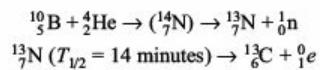
They bombarded aluminium with α -particles from a polonium source. They found that even after the polonium source was removed, the target nuclei continued to eject positrons (positron is a particle having mass equal to that of an electron and charge equal to that of a proton, discovered by Anderson in 1932). This can be represented by



The phosphorus nucleus so produced is not stable but radioactive. The radiophosphorus ($T_{1/2} = 3.25$ minute) disintegrates spontaneously by emitting a positron and neutrino and finally changes into a stable silicon atom. This is represented below.



Similar reaction with a boron nucleus is as follows.



At present, more than 1300 artificial radionuclides are known to exist. The study of artificial radioactivity has helped in clarifying the concepts regarding the constitution and stability of the nuclei.

The transuranium elements ($Z > 92$) such as ${}_{93}\text{Np}$, ${}_{94}\text{Pu}$, ${}_{95}\text{Am}$, ${}_{96}\text{Cm}$, etc. are not naturally occurring elements and they are produced in the laboratory using artificial radioactivity.

3.9 APPLICATIONS OF RADIOACTIVITY

We discuss below few important applications of radioactivity.

- (i) **Dating:** To estimate the age of archaeological and geological specimens, the exponential decay law of radioactive nuclei is used. This technique is known as *dating*.

The half-life of ${}^{14}_6\text{C}$ nucleus is 5730 years. This radioactive isotope of carbon is being created continuously in the atmosphere by high energy particles coming from outer space (called cosmic rays), as per the following reaction



That is, when a nitrogen atom is hit by a neutron, the radioactive isotope ${}^{14}_6\text{C}$ and ${}^1\text{H}$ is created. Eventually, kind of equilibrium has been established between the production and decay of ${}^{14}_6\text{C}$ atoms. Therefore, in the atmosphere there exists a small fraction of ${}^{14}_6\text{C}$ atoms in the total CO_2 molecules. The living organisms utilize carbon from the atmosphere in its metabolic cycle, a certain fraction of which is ${}^{14}_6\text{C}$ atoms. When the organisms die, its intake of ${}^{14}_6\text{C}$ stops and the ${}^{14}_6\text{C}$ atoms already present starts decaying. The age of an archaeological specimen can be determined by the amount of ${}^{14}_6\text{C}$ it contains and comparing this amount with the content of ${}^{14}_6\text{C}$ in a live similar organism. Similarly, dating using uranium content in a geological specimen can also be carried out.

- (ii) **Tracer Technique:** The chemical properties of radioisotopes of an element are identical with its stable isotope. Therefore, if a mixture of stable isotope and radioisotope of an element (for example, iodine with radioisotope ${}^{131}\text{I}$ with $T_{1/2} = 12.6$ hours) is introduced into a system, both the isotopes will undergo the same chemical, biological and physical processes, except the radioactivity shown by the radioisotope. The radioisotope iodine decays via β -decay (e^- emission). Therefore, using an electron detector the movement of the radioisotope can be traced. This technique is called tracer technique often used in medical diagnostics and also for certain industrial applications. Doctors use tracer technique to locate the blockage in blood flow before operating for the removal of blockage.

It has been found that cancerous tissues retain larger amount of bismuth than normal tissues. By administrating a mixture of ordinary bismuth with a small amount of radioactive bismuth ${}^{210}\text{Bi}$ to a person, the cancerous tissues can be located by testing the radioactivity over various parts of the body.

This technique can also be used to find out the wear and tear. For example, the wear of a piston ring can be found out by incorporating in it a radioisotope and then measuring the

accumulation of activity in the lubricating oil. Similarly, the wear of a car type can be investigated by incorporating radioisotope of phosphorus ^{32}P into its outer tread and observing the decrease in counting rate as a function of time.

- (iii) **Therapeutic Use of Radioactivity:** Exposure to radiations emitted by radioactive elements is harmful for living cells and an excessive exposure can even cause death. However, this effect has been used in cancer treatment. A controlled administration of radiation from radioisotopes is used to cure or contain cancer. Radioactive isotope of iodine ^{131}I has been used to reduce enlarged thyroid to its normal size.
- (iv) **Activation Analysis:** In this technique, the sample to be analyzed is bombarded with neutrons. The addition of a neutron to the nucleus of an element makes it unstable and induce radioactivity. By analyzing the radiation emitted by such nuclei, the element can be identified. This technique is particularly used for identifying the elements occurring in traces. For example, in a sample of a material containing traces of gold ^{197}Au , the concentration of gold can be determined by bombarding the sample with neutrons and measuring the activity of ^{197}Au . The process is repeated with a sample containing a known amount of ^{197}Au . Using this as standard, the amount of ^{197}Au in the sample can be estimated. This method can also be estimated for the presence of an element in a patient's body.
- (v) **As Preservative:** Since the nuclear radiations kill microorganisms, exposure to the nuclear radiation of foods, increase its shelf life. That is, the process can be used for food preservation. It has been found that irradiated potatoes and onions remain free from sprouting diseases for a longer period of time.
- (vi) **Agriculture:** Radioisotopes are widely used in agriculture research. For example, radioisotope ^{13}N (β -emitter) is used to investigate the uptake of nitrogen by plant roots and the movement of nitrogen at the growth points. Irradiated seeds remains free from the germs for a longer period of time.

3.10 ALPHA DECAY

The nuclei with $Z > 82$ are unstable and tend to transform into more stable configuration by emitting nuclear radiations. As discussed earlier, each atomic nucleus has a neutron-proton ratio in a certain range for which nuclei are stable. All the nuclei having this ratio outside stability range are radioactive and decay in such a way that it brings back the neutron-proton ratio within the stability range. Radioisotopes, both natural and artificial, emit α -particles (^4_2He nucleus), β -particles (electrons and positrons) and γ -rays (photons) thereby reducing their energy and achieving stability. Starting with α -decay, we discuss below the three modes of decay of unstable nuclei.

In α -decay an unstable nucleus disintegrates into a daughter nucleus whose atomic number is reduced by 2 and atomic mass number is reduced by 4. It is represented as



The quantity Q is the energy released in the decay process and is given by

$$Q = [M_x - (M_y + M_\alpha)] \cdot c^2 \quad (3.19)$$

where, M_x , M_y and M_α are the masses of the parent nucleus, daughter nucleus and α -particle respectively. In such decays, the total mass of the decay product is less than the mass of the original nucleus.

The kinetic energy of the released α -particle can be determined by applying the laws of conservation of energy and momentum. Applying the law of conservation of energy

$$[M_x - (M_y + M_\alpha)] \times c^2 = \frac{1}{2} M_\alpha v_\alpha^2 + \frac{1}{2} M_y v_y^2 \quad (3.20)$$

where v_α and v_y are the velocities of the α -particle and the daughter nucleus. Since the parent nucleus is initially at rest, its initial momentum is zero. Using the law of conservation of momentum, we have

$$M_\alpha v_\alpha = M_y v_y \quad (3.21)$$

Substituting the value of v_y from Eq. (3.21) in Eq. (3.20), we get

$$\begin{aligned} Q &= \frac{1}{2} M_\alpha v_\alpha^2 + \frac{M_y M_\alpha^2 v_\alpha^2}{2 M_y^2} \\ &= \frac{1}{2} M_\alpha v_\alpha^2 \left(1 + \frac{M_\alpha}{M_y}\right) \\ &= K_\alpha \left(\frac{M_\alpha}{M_y} + 1\right) \quad (K_\alpha = \text{kinetic energy of } \alpha\text{-particle}) \end{aligned}$$

Since, $\frac{M_\alpha}{M_y} = \frac{4}{A-4} \Rightarrow \frac{M_\alpha}{M_y} + 1 \approx \frac{4}{A-4}$ ($A \gg 4$)

$$\therefore K_\alpha \approx \frac{A-4}{4} Q \quad (3.22)$$

As $A \gg 4$, $\frac{A-4}{4} \approx 1$. This implies $K_\alpha \approx Q$, i.e., almost the entire energy released in the decay appears as the kinetic energy of the α -particle.

As an example, let us consider the α -decay of $^{226}_{88}\text{Ra}$, i.e.,



In this case, the value of Q is found to be 4.87 MeV and the emitted α -particle is observed to have an energy 4.78 MeV. In the above example, the α -particle has a definite energy because both the parent and daughter nuclei are in a definite energy state. However, in some cases the α -particle spectrum is complex and includes many α -particles with different energies. This is because the daughter nucleus can be in its various possible excited states.

Table 3.2 shows the half-lives of 5 different α -radioactive nuclei. It can be surprisingly observed that the half-life varies from billions of years to a fraction of microsecond and energy released range only from about 4 MeV to 9 MeV. Also, it is worth noting that *shorter the half-life, the higher is the energy released and vice versa*.

Table. 3.2 Kinetic energy released and half-lives of five α -radioactive nuclei

Nucleus	$K(\text{MeV})$	$T_{1/2}$
^{232}Th	4.1	14 billion years
^{226}Ra	4.9	1600 years
^{240}Cm	6.4	27 days
^{194}Po	7.0	0.7 second
^{216}Ra	9.5	0.18 μs

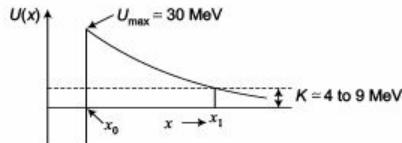
Theory of α -decay: The underlying assumption of the theory of α -decay is that the two of the protons inside a heavy nucleus can occasionally cluster together with two of the neutrons to form an α -particle. Experiments have shown that this is a correct assumption. Alpha particles are continually forming and dissolving inside a nucleus. Averaged over time, it is estimated that there is one alpha particle inside any nucleus.

An α -particle is formed from four nucleons, each with potential energy of about -50 MeV and kinetic energy of about 25 MeV . Therefore, an α -particle should have an energy of about 100 MeV which corresponds to a speed $v \approx \frac{c}{4}$. For a nuclear diameter of about 15 fm , the frequency with

which an α -particle appears at the surface of the nucleus around $f \approx \frac{v}{D} \approx \frac{c}{4 \times 15 \text{ fm}} \approx 5 \times 10^{21} \text{ s}^{-1}$.

That is, α -particles appear at the nuclear surface 10^{21} times per second. However, these particles mostly do not escape the nucleus since the strong nuclear force pulls them back into the nucleus. But occasionally an α -particle escape a nucleus.

The theory of α -decay was worked out by George Gamow and others on the basis of quantum tunnelling. In Gamow's theory, α -particle is assumed to be formed and bound inside the nucleus by nuclear forces. The potential energy of α -particle is shown in Fig. 3.7. If the α -particle is outside the nucleus, then U is just the coulomb potential energy $U(x) = \frac{2 Z k e^2}{x}$, where x is the separation between the α -particle and the residual nucleus and $2e$ and ze are the charges on the α -particle and the residual nucleus respectively. For x less than x_0 (\approx nuclear radius R), the attractive nuclear force dominates and the potential energy drops abruptly. There is thus a barrier and this is the barrier which prevents the α -particle from escaping instantly. However, in quantum mechanics tunnelling of a particle through a barrier is possible which allows the α -particle to pass through the barrier and escape.

**Fig. 3.7** Potential energy of α -particle which confine it within nucleus.

To calculate the probability of such tunnelling, we need to establish the approximate parameters of the barrier confining the α -particle. Its left hand boundary is at the point x_0 , which is about equal to the nuclear radius (or more precisely, the sum of the radii of the residual nucleus and the α -particle):

$$x_0 \approx 8 \text{ fm}$$

The height of the barrier is equal to the electrostatic potential energy at $x = x_0$ (as shown in Fig. 3.7):

$$U_{\max} = \frac{2 Z k e^2}{x_0} \sim 2 \times 90 \times \frac{1.44 \text{ MeV}}{8 \text{ fm}} \approx 30 \text{ MeV}$$

We know that when the α -particle emerges from the nucleus and moves far away, its kinetic energy (or total energy since $U = 0$) lies in the range of 4 to 9 MeV, shown by the dashed line in Fig. 3.7. This is much below the top of the barrier ($U_{\max} \approx 30$). Therefore, from the classical considerations the α -particle should be permanently trapped by the barrier.

Another important parameter shown in Fig. 3.7 is the distance x_1 . This is the separation at which the particles coming towards the nucleus with initial kinetic energy K would be stopped by the barrier. The value of x_1 is determined by the condition

$$\frac{2Z k e^2}{x_1} = K$$

$$\text{or } x_1 = \frac{2Z k e^2}{K} \approx 2 \times 90 \times \frac{1.44 \text{ MeV.fm}}{4 \text{ to } 9 \text{ MeV}} \approx 30 \text{ to } 65 \text{ fm}$$

The interval from x_0 to x_1 is the classically forbidden region. Its length, $L = x_1 - x_0$ is the thickness of the barrier through which the α -particle must tunnel through for the decay to occur. The value of L varies from 20 to 55 fm as K ranges from 9 MeV to 4 MeV, which is extremely long by the nuclear standards. These are why the escape probability P is very low and why it varies so much with the energy of the emerging α -particle.

Probability of Tunnelling. We know that in quantum mechanics it is possible that a particle tunnel through a potential barrier which is classically not possible. For a simple rectangular barrier (Fig. 3.8), we have seen that the probability of a particle with energy $E < U_0$ tunnelling through a barrier of length L is given by

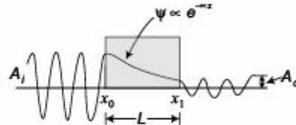


Fig. 3.8 A rectangular potential barrier and ψ across it for, $x < x_0$, ψ is sinusoidal, within the barrier it decreases exponentially, for $x > x_1$, it is again sinusoidal with amplitude A_o

$$P \approx e^{-2\alpha L} \quad (3.24)$$

where

$$\alpha = \sqrt{\frac{2m(U_0 - E)}{\hbar^2}} \quad (3.25)$$

and m is the mass of the particle whose tunnelling is being considered.

The tunnelling probability depends on the two factors: α given by Eq. (3.25) and L , the length of the barrier. If the barrier is high and wide, as in the case of alpha-decay, both α and L are large and the probability is very small (as if x is large, e^{-x} is extremely small). Further, if α is large, $P = e^{-2\alpha L}$ is very sensitive to the barrier thickness L . This is the main reason why half-lives vary over such an enormous range.

Equation (3.24) gives the tunnelling probability for a particle confined by a rectangular barrier. For an alpha particle confined by a non-rectangular barrier of Fig. 3.7 to find the tunnelling probability, we approximate the actual barrier by a succession of n rectangular barriers as shown in Fig. 3.9. The total tunnelling probability is the product of the n individual probabilities

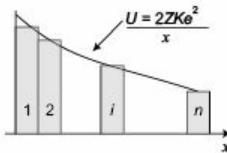


Fig. 3.9 Approximation of barrier.

$$\begin{aligned} P &= P_1 P_2 \dots P_n \\ &= e^{-2\alpha_1 \Delta x_1}, e^{-2\alpha_2 \Delta x_2} \dots e^{-2\alpha_n \Delta x_n} \\ &= \exp (-2 \sum \alpha_i \Delta x_i) \end{aligned} \quad (3.26)$$

where each α_i given by Eq. (3.25) appropriate barrier height U_i . In the limit, $\Delta x_i \rightarrow 0$, the summation becomes an integral and we can express the probability as

$$P = \exp \left[-2 \int_{x_0}^x \alpha(x) dx \right] \quad (3.27)$$

where $\alpha(x) = \sqrt{\frac{2m[U(x) - K]}{\hbar^2}}$ (3.28)

by replacing U_0 with $U(x)$ and noting that E is just K , the kinetic energy when the alpha particle is far away from the nucleus and U is zero.

For alpha-decay, $U(x) = \frac{2Zke^2}{x}$, and the integral in Eq. (3.27) can be evaluated to give

$$P = \exp \left[\frac{-2\pi ke^2 \sqrt{2m}}{\hbar} Z K^{-1/2} + \frac{8\sqrt{mke^2}}{\hbar} (ZR)^{1/2} \right] \quad (3.29)$$

where Z and R denote the charge and radius of the residual nucleus. It is convenient to express this as

$$P = \exp [-aZ K^{-1/2} + b(ZR)^{1/2}] \quad (3.30)$$

where the constants a and b are given by

$$a = \frac{2\pi ke^2 \sqrt{2m}}{\hbar} = 3.97 \text{ (MeV)}^{1/2} \quad (3.31)$$

and $b = \frac{8\sqrt{mk e^2}}{\hbar} = 2.98 \text{ (fm)}^{-1/2}$ (3.32)

Using $Z = 90$, $K = 6 \text{ MeV}$ and $R = 8 \text{ fm}$ into Eq. (3.30), we find

$$P \approx \exp(-146 + 80) \approx 2 \times 10^{-29}$$

Obviously, the probability that an alpha-particle striking the surface of a nucleus will tunnel through is extremely small.

In terms of half-lives: Experimental data is generally expressed in terms of half-lives, we express our results in terms of the half-life of an alpha particle emitting nucleus. We know, $T_{1/2} = \frac{\ln 2}{\lambda}$ and the decay constant $\lambda = fP$, where f is the frequency with which alpha particles strike the nuclear surface.

Therefore,

$$T_{1/2} = \frac{\ln 2}{\lambda} = \frac{0.693}{fP}$$

Since P is in exponential form, it is convenient to take natural logarithm of the expression, to obtain

$$\ln T_{1/2} = -\ln P + \ln \frac{0.693}{f}$$

Substituting for P , we get

$$\ln T_{1/2} = a ZK^{-1/2} - b (ZR)^{1/2} + c \quad (3.33)$$

where Z is the charge and R the radius of the residual nucleus. The constants a and b are given by Eqs. (3.31) and (3.32) respectively and the constant c is $\ln \frac{0.693}{f}$, where f is approximately $5 \times 10^{21} \text{ sec}^{-1}$, therefore

$$c = \ln \frac{0.693}{f} \approx -50$$

Half-lives of nuclei vary because K , Z and R are variables. The singlemost important factor is the energy dependence of the term $a ZK^{-1/2}$. Small differences in K correspond to appreciable difference in this term and hence enormous difference in $\ln T_{1/2}$, particularly as K vary from 4 to 9 MeV, $T_{1/2}$ drops by 24 orders of magnitude.

3.11 BETA DECAY

In beta-decay, a nucleus ejects an electron (β -particle) or a positron (β^+ – particle) or an orbital electron is absorbed by the nucleus. In this, the atomic number of the nucleus changes by one unit but the mass number remains unchanged. In other words, the charge of the nucleus changes but the number of nucleons remains the same. Since a nucleus contains only the neutrons and protons,

how the emission of an electron or a positron can be explained? This could be explained by assuming that these particles are produced during the emission process (formed at the instant of emission—cannot reside inside the nucleus) in the same way a photon is produced during an emission from an excited state. Within a nucleus, a neutron changes to a proton and an electron



The Q -value of this process is given by

$$\begin{aligned} Q &= (m_n - m_p - m_e) \cdot c^2 = [m_n - m (\text{H})] \cdot c^2 \\ &= (1.008665 u - 1.007825 u) 931.5 \text{ MeV/u} = 0.728 \text{ MeV} \end{aligned}$$

Since the recoil energy of the proton is very small, it is expected that this energy should appear as the kinetic energy of the emitted electron. In other words, all the emitted electrons should have exactly this energy. However, it was found experimentally that the emitted β -particles have a continuous range of energy varying from zero to a maximum (≈ 0.728 MeV as calculated above). Initially, this observation proved to be quite a problem and different approaches were tried to explain this observation. Even the possibility of the violation of the law of conservation was considered. Detailed experiments with decay indicated that conservation of angular momentum was doubtful. The decay expressed in Eq. (3.34) also violates this clearly. The spin angular momentum of neutron

is $\frac{\hbar}{2}$ while the proton and electron, each having a spin $\frac{1}{2}$ should give a total angular momentum of 0 or $1\hbar$.

The violation of the two fundamental conservation laws—the law of conservation of energy and angular momentum, were troubling the physicists. Pauli in 1930 made a pathbreaking postulate. He postulated that a third particle was emitted in the beta-decay process. He argued that as the electric charge is already conserved by the proton and electron, the third particle must be neutral one. The missing energy is carried by this particle. This particle should have a spin $\frac{1}{2}$, because with 0 or 1 the

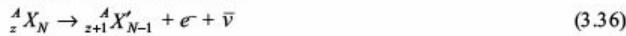
proton-electron combination, the resultant can be of spin $\frac{1}{2}$, as required by the law of conservation of angular momentum. He proposed that particle should have zero rest mass. This particle was named *neutrino* (ν) by E. Fermi. Later the existence of this particle was confirmed experimentally.

However, the difficulty of observing the neutrino is that it interacts very weakly with the matter. In light of this particle, we discuss below the beta-decay process.

Electron Emission: The antiparticle of neutrino ν is the antineutrino ($\bar{\nu}$). The decay in Eq. (3.34) can be expressed as



In general, electron emission can be expressed as



The Q -value for this decay is

$$Q = [m\left(\frac{A}{z}X\right) - m\left(\frac{A}{z+1}X'\right)]c^2 \quad (3.38)$$

only the atomic masses appear in the above equation as the electron masses cancel out
 $[Q = [m\left(\frac{A}{z}X\right) - Zm_e - m\left(\frac{A}{z+1}X'\right) - (Z + 1)m_e + m_e]c^2]$. The rest mass of the neutrino (or antineutrino) is assumed to be zero and so does not appear in the equation. Since the kinetic energy of the recoil nucleus (proton) is negligibly small, the energy released in the decay process appears as the kinetic energy of the electron and the antineutrino particle. That is,

$$Q \approx K_e + E_{\bar{\nu}}$$

The energy of the emitted electrons is not discrete, unlike in α -decay, but continuous, as shown in Fig. 3.10. This continuous nature of energy distribution posed quite a problem in understanding the beta-decay process. This problem was resolved after the prediction of the existence of the neutrino which was subsequently discovered in 1956.

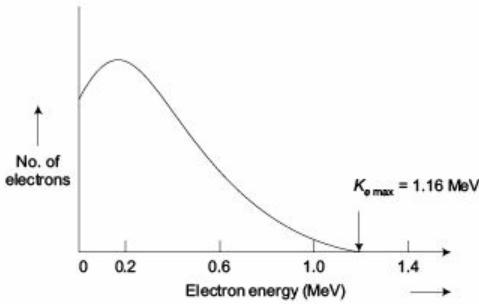


Fig. 3.10 Energy distribution in β^- -decay.

The beta decay results in three particles — a beta particle, a neutrino (or antineutrino) and a daughter nucleus. The neutrino carries some energy and so the energy of the electron is less than what it would have been if only the daughter nucleus and the beta-particle were produced in the decay. Based on neutrino hypothesis, Fermi could explain the β -decay process. Depending upon the amount of energy carried by neutrino, the energy of the emitted particle vary from zero to a maximum (≈ 1.16 MeV).

The neutron and the proton that are transformed are not free but bound to the nucleus by nuclear forces. According to Fermi, an interaction that exists between the neutron, electron and antineutrino or the proton, positron and neutrino, is responsible for the transformation of a neutron into a proton or proton into a neutron. It is a very weak interaction. Thus, the interaction called *weak-interaction* is responsible for β -decay. Using half-life of the β -decay, the magnitude of the β -decay interaction was estimated to be of the order of 10^{-13} times weaker than the strong interaction.

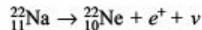
Positron Emission: In some β -decay process, positron (β^+ -particle) is ejected. Positron is the antiparticle of the electron. It has the same mass as that of an electron but a positive charge $+ e$. The decay can be represented as



where e^+ represents the positron and ν the neutrino (even though ν and $\bar{\nu}$ are different particles—on many occasions the term ‘neutrino’ is used for both the particles where the distinction is not important). A typical β^- -decay can be represented as



For example,

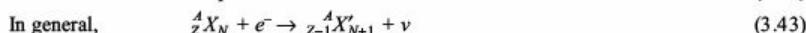


The Q -value of this decay is given by

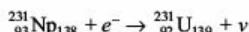
$$Q = [m({}_{Z}^A X) - m({}_{z-1}^{A'} X'_{N+1})] - 2m_e c^2 \quad (3.41)$$

For the positron emission to occur, Q must be positive, i.e., the initial atomic mass must exceed the final atomic mass. The positron and the neutrino share this energy. The energy distribution of the positrons vary continuously as in the case of electron emission.

Electron Capture. In addition to β^- and β^+ decays, sometimes a proton in a nucleus captures one of its atomic electrons and converts itself into a neutron and a neutrino



For example,



Usually, the electron that is captured is one of the K-shell electrons, therefore, the process is sometimes referred as **K-capture**. If the initial kinetic energy of the electron and the recoil energy of the nucleus are neglected, the entire energy Q is taken away by the neutrino.

Example 3.11

Find the kinetic energy of the electrons ejected in the β^- -decay of bismuth 210 to polonium 210. Given $m({}^{210}\text{Bi}) = 209.984096 \text{ u}$ and $m({}^{210}\text{Po}) = 209.982848 \text{ u}$.

Solution. We know, the electrons kinetic energy is given by

$$K = \Delta m \cdot c^2$$

Here, $\Delta m = 209.984096 - 209.982848 = 0.001248 \text{ u}$

$$\therefore \text{Kinetic energy, } K = 0.0012484 \times 931.5 \frac{\text{MeV}/c^2}{u} = 1.16 \text{ MeV}$$

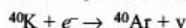
If all the energy is taken by the electron.

Example 3.12

${}^{40}\text{K}$ decays to ${}^{40}\text{Ar}$ by electron capture. Calculate the kinetic energy of the neutrino assuming that the initial kinetic energy of the electron and the recoil energy of the nucleus are zero.

Given: Mass of ${}^{40}\text{K} = 39.96399 \text{ u}$; mass of ${}^{40}\text{Ar} = 39.962384 \text{ u}$

Solution. The electron capture can be represented as



$$\therefore Q = [m({}^{40}\text{K}) - m({}^{40}\text{Ar})] \cdot c^2$$

$$\begin{aligned}
 &= [39.96399 u - 39.962384 u] \times 931.5 \frac{\text{MeV}/c^2}{u} \\
 &= 1.504 \text{ MeV}
 \end{aligned}$$

Thus, the kinetic energy of the neutrino is 1.504 MeV.

3.12 GAMMA DECAY

A nucleus, like an atom, exists in one of a number of quantized energy states. A stable nucleus, is normally in its lowest (i.e., ground) energy state. If a nucleus, is in excited state, it decays into one of its lower states, releasing the energy difference in the form of electromagnetic radiation. After alpha or beta decay, a nucleus may be left in an excited state. The separation between the nuclear energy levels lies between a few tens of keV to a few MeV, unlike a few eV in atoms. Hence, the energy of the photons given out lie in the range of keV to MeV. The radiation so emitted is called *gamma-ray* and the process is termed as *gamma-decay*. Through this process a nucleus attain ground state directly or a lower excited state which in turn decays to a still lower excited state or ground state, emitting in this process a cascade of photons. In this process, however, only those nuclear transitions are allowed which obey the laws of conservation of energy, momentum, angular momentum and parity. In other words, nuclear transitions are also governed by selection rules applicable to the atomic transitions. An excited state of a nucleus has a very short half-life of the order of 10^{-14} to 10^{-9} sec. However, in certain cases it may extend to several years. Such delayed transitions are called *isomeric transitions* and such excited states are called *isomeric states* or *isomeric levels*.

Nuclei having the same Z and A but different radioactive properties are called *nuclear isomers* and the phenomenon is called *nuclear isomerism*.

Figure 3.11 shows gamma-decay in $^{60}_{27}\text{Co}$ nucleus. For this nucleus, $T_{1/2} = 5.2$ years. A beta-decay leaves the daughter nucleus $^{60}_{28}\text{Ni}$ into an excited state which emits two gamma-rays successively with energies 1.17 MeV and 1.33 MeV. The $^{60}_{27}\text{Co}$ is commonly used as the source of gamma-rays.

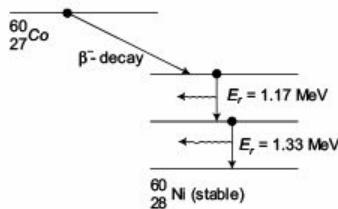


Fig. 3.11 Gamma-decay in $^{60}_{27}\text{Co}$ source.

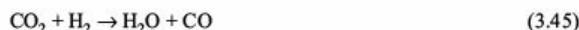
3.13 NUCLEAR REACTIONS

Certain nuclear transformations occur spontaneously, i.e., unstable nuclei decay without any external stimulus, resulting in a new nucleus. There is another kind of nuclear transformation, in which a target nucleus is bombarded by fast moving particle such as protons, neutrons, deuterons, α -particles, etc. The interaction between the target nucleus and the incident particles depends on the nature and energy of the particles and the target nucleus. The mass number or the atomic number or both of the target nucleus may change and a completely different nucleus may be produced. Such transformation of a nucleus into another one is called *transmutation* while the process is termed *nuclear reaction*. The first nuclear reaction was observed by Rutherford in 1919 when he bombarded nitrogen nuclei with α -particles and produced the following reaction.



The above equation is also expressed as ${}_{14}^1\text{N}(\alpha, p){}_{8}^{17}\text{O}$ in short. Since then, innumerable number of nuclear reactions have been reported. Very often, the energy of the projectile determines the outcome of the nuclear reaction. Sometimes, the product nucleus is radioactive which undergo a series of radioactive decay before achieving stability.

The reason for calling the transformation, a nuclear reaction is its resemblance to any chemical reaction such as



in which molecules (or atoms) collide and rearrange the constituents to give new products. However, there is a major difference between chemical reactions and nuclear reactions, while the chemical reactions occur naturally and abundantly, hardly any nuclear reactions takes place in the normal circumstances on the earth. The reason for this is as follows. For two atoms or molecules to react chemically, they only need to come to each other so that their outer electrons overlap which is easily possible at normal temperature and pressure. However, nuclei being positively charged, the Coulomb repulsion prevent them to come close enough for the nuclear reactions to occur, at normal temperatures and densities. However, in stars, the temperature is so high that the kinetic energy of thermal motion is sufficient to overcome the Coulomb repulsion, and the nuclear reactions do occur.

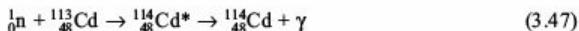
To produce nuclear reactions in laboratory, the energies of the nuclei should be increased to order of 1 MeV or more and then they should hit target nuclei. The energy of the projectiles, enables some of them to approach the target nucleus close enough to induce a nuclear reaction. Till 1930s, the only available projectiles were the α -particles from natural radioactive sources. After the development of particle accelerator, it became possible to accelerate many types of nuclei to an energy sufficient to induce nuclear reactions. Also, discovery of neutron made it possible to induce nuclear reactions using even low energy neutron, as neutron being neutral can approach and penetrate a nucleus. This method has become quite popular in nuclear research.

An example of a nuclear reaction induced by a projectile accelerated by an accelerator is given below.



In this reaction, accelerated protons were used to hit ${}^7\text{Li}$ target.

An example of a neutron induced nuclear reaction is given below.



When a neutron strike ${}_{48}^{113}\text{Cd}$, a compound nucleus ${}_{48}^{114}\text{Cd}^*$ is formed, which subsequently emit a γ -ray photon.

In a nuclear reaction, number of protons, number of neutrons, energy, linear momentum and angular momentum are conserved. In a nuclear reaction, the projectile and target nucleus are within the range of one another's influence only for extremely short time of about 10^{-20} sec.

Reaction Cross-section. All incident projectile particles do not produce a nuclear reaction, only a small fraction of them produce a nuclear reaction. For better understanding of nuclear reactions, it is desirable to know the probability that a single projectile particle will interact with an individual target nucleus. Such a probability is expressed in terms of a parameter called the *reaction cross-section*, which is defined as follows.

Suppose a beam of projectiles approaches a single target nucleus, such that n *number of projectiles are incident per unit area perpendicular to the incident direction*. That is, n is *flux of projectiles*, expressed in particles/area. Now, Let N be *number of reactions* that occur due to the incident flux. Then, the *reaction cross-section* (or simply *cross-section*) σ is expressed as

$$\sigma = \frac{N}{n} = \frac{\text{Number of reactions}}{\text{Flux of projectiles}} \quad (3.48)$$

Obviously, the cross-section σ is large for reactions that are highly probable and small for those which are unlikely to occur. Since σ has the dimensions of area, we can imagine the cross-section to the *size as seen by the projectile*. Therefore, in a sense, the cross-section σ is the cross-sectional area of the target and in a nucleus it represents the effective area of a target for reacting with a particular projectile. For example, the cross-section of inverse β -decay



is found to be about $6 \times 10^{-48} \text{ m}^2$ (very small) indicating that as seen by a neutrino, the proton is exceedingly small and so almost all neutrinos pass unaffected. The inverse β -decay process is highly probable.

For most nuclear reactions, the cross-section is found to be of the order of the geometrical size of nuclei. Since the typical nuclear size is of the order of 10^{-14} m , the unit of area of cross-section is 10^{-28} m^2 , expressed in barn. That is,

$$1 \text{ barn} = 10^{-28} \text{ m}^2.$$

The value of σ also depends upon the energy of the projectiles. Knowing σ , we can estimate the number of reactions that will occur. From Eq. (3.46), the number of reactions produced by a single target nucleus is

$$N \text{ (for a single target nucleus)} = \sigma n \quad (3.50)$$

In practice, all targets contain many nuclei. If N_{tar} is the total number of target nuclei, then total number of reactions is given by

$$N = \sigma n N_{\text{tar}} \quad (3.51)$$

3.13.1 Q Value of a Nuclear Reaction

Consider a nuclear reaction represented as



Let the rest mass of the incident projectile particle be m_x and its kinetic energy be K_x , while that of target nucleus be m_X and K_X respectively. Further, let the rest mass of outgoing particle and the recoil nucleus be m_y and M_Y and their kinetic energies be K_y and K_Y respectively. The nuclear reaction is depicted in Fig. 3.12. Along with the other conservation laws, the law of conservation of relativistic energy apply to all nuclear reactions. Applying this to the above-mentioned nuclear reaction, we have

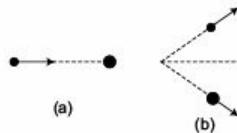


Fig. 3.12 Illustration of a nuclear reaction, (a) before and (b) after collision.

$$K_x + m_x c^2 + K_X + M_X c^2 = K_y + m_y c^2 + K_Y + M_Y c^2 \quad (3.53)$$

The masses in the above equation refers to the nuclear rest masses. Rearranging the above equation, we get

$$(K_y + K_Y) - (K_x + K_X) = (m_x + M_X) c^2 - (m_y + M_Y) c^2 \quad (3.54)$$

$$\text{or } \Delta K \equiv Q = K_f - K_i = (M_i - M_f) c^2 \quad (3.55)$$

where $K_x + K_X = K_p$ the total initial kinetic energy (normally $K_X = 0$ as the target nucleus is at rest) and $K_y + K_Y = K_f$ is the total final energy. Similarly, the initial mass $M_i = m_x + M_X$ and the final mass $M_f = m_y + M_Y$. From Eq. (3.55), it is implied that that the net increase in kinetic energy is equal to the net decrease in rest mass. *The gain in kinetic energy in a nuclear reaction is called the Q-Value of the nuclear reaction* and it is denoted by Q . The calculation of ΔK , provides a way for the experimental verifications of Einstein's mass-energy relation.

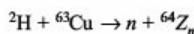
If $Q > 0$, the reaction is said to be *exothermic* (or *exoergic*) as the energy is released in the reaction and therefore the total kinetic energy after the reaction is greater than that before the reaction. Thus, $Q > 0$ implies $K_f > K_i$ or $M_i > M_f$. If $Q < 0$, the reaction is *endothermic* (or *endoergic*) as the energy is required for the reaction to occur in the form of kinetic energy of x which should provide the additional rest energy of the reaction products. Therefore, the kinetic energy of the projectile must have some minimum or threshold value below which the reaction cannot take place. For the products to have some kinetic energy, the kinetic energy of x must exceed the threshold value. Thus,

$$Q < 0 \text{ implies } K_f < K_i \text{ or } M_i < M_f.$$

Example 3.13

Deutrons of energy 10 MeV are incident on a ^{63}Cu target and neutrons are observed with 15 MeV of kinetic energy. Evaluate (i) the Q-Value for the reaction, (ii) the kinetic energy of the residual nucleus ^{64}Zn . The atomic masses are: ${}^2\text{H} = 2.014102 \text{ u}$, $n = 1.008665 \text{ u}$, ${}^{63}\text{Cu} = 62.929599 \text{ u}$ and ${}^{64}\text{Zn} = 63.929144 \text{ u}$.

Solution. The involved nuclear reaction is



- (i) The Q -value of the reaction is

$$Q = (2.014102 \text{ } u + 62.929599 \text{ } u - 1.008665 \text{ } u - 63.929144) \times 931.5 \frac{\text{MeV}}{u}$$

$$= 5.488 \text{ MeV}$$

- (ii) The kinetic energy of the residual nucleus is

$$K_T = Q - K_x - K_y = (5.488 + 10 - 15) \text{ MeV} = 0.488 \text{ MeV.}$$

3.14 NUCLEAR MODELS

We often try to understand something not well understood in terms of something well understood, i.e., in terms of certain existing 'models'. In early days of the development of the nuclear physics, many models of the nucleus were proposed to explain the observed properties of the nucleus. Two of these models – the liquid drop model and the shell model are discussed below.

3.14.1 Liquid Drop Model

The liquid drop model of a nucleus is based on the two properties possessed by all nuclei, except those with very small A . These are: (i) the density of the nuclear matter is independent of the size of the nucleus and (ii) the binding energy per nucleon is nearly constant.

Both these properties are analogous to those found in a drop of liquid. Therefore, in this model a nucleus is considered to be analogous to a liquid drop. It is well known that to convert a liquid into vapour, the energy required is directly proportional to the mass of the liquid. This means that the total binding energy of all the molecules in the liquid is directly proportional to the mass of the liquid. In other words, the binding energy per unit mass is constant. Further, the density of a liquid drop is constant and does not depend upon the size or shape of drop. Thus, both these properties are similar in the liquid drop and the nucleus.

Weizsäcker Semiempirical Formula for Binding Energy

The liquid drop model approximates the nucleus as a sphere with uniform density. Most of the observed properties of nuclei could be explained satisfactorily using this model. The radius is proportional to $A^{1/3}$, the surface area proportional to $A^{2/3}$ (area = $4\pi r^2$), and the volume is proportional to A ($\text{volume} = \frac{4}{3}\pi r^3$). Since the mass is A , the density ($= \frac{\text{mass}}{\text{volume}}$) is a constant. The shape of

the nucleus is determined mainly by three factors, the volume energy, the surface energy and the Coulomb energy. The analogy of a nucleus with a liquid drop was proposed by George Gamow in 1929 and developed in detail by C.F. Von Weizsäcker in 1935.

Volume energy. Assume energy associated with each nucleon-nucleon bond has a value U (although U should have a negative sign as the force involved is attractive but it is usually written as

positive since the binding energy is considered positive quantity for convenience). Because each bond energy U is shared by two nucleons, each has a binding energy $\frac{1}{2}U$. When an assembly of spheres of the same size is packed together into small volume, as should be the case with the nucleons within a nucleus, each interior sphere has 12 other spheres in contact with it. Hence, each interior nucleon in a nucleus has a binding energy of $12 \times \frac{1}{2}U = 6U$. If all A nucleons in a nucleus were in its interior, the total binding energy of the nucleus would be

$$E_v = 6A U \quad (3.56)$$

The above equation is often expressed as

$$E_v = a_1 A \quad (3.57)$$

where a_1 , is a constant.

The energy E_v is called the *volume energy* of a nucleus and is directly proportional to A .

Surface Energy. Some nucleolus are on the surface of the nucleus and therefore have less than 12 neighbouring nucleons. The number of such nucleons depends on the surface area of the nucleus. A nucleus of radius R has an area $4\pi R^2 = 4\pi R_o^2 A^{2/3}$. Hence, the number of nucleons with less than the maximum number of bonds is proportional to $A^{2/3}$, reducing the binding energy by

$$E_s = -a_2 A^{2/3} \quad (3.58)$$

where a_2 is a constant.

The negative energy E_s is called the *surface energy* of a nucleus. It is very significant for the lighter nuclei since a greater fraction of their nucleons are on the surface. Because energy system tends to attain minimum potential energy, nuclei also tend to attain configurations of maximum binding energy. Hence, a nucleus should exhibit the same surface tension effects as a liquid drop and in the absence of other effects it should be spherical, since a sphere has the least surface area for a given volume.

Coulomb Energy. The electrostatic repulsion between each pair of protons in a nucleus also contributes toward decreasing its binding energy. The *coulomb energy* E_c of a nucleus is the work that must be done to bring together Z protons from infinity into a spherical aggregate of the size of the nucleus. The potential energy of a pair of protons separated by r is equal to

$$V = -\frac{e^2}{4\pi\epsilon_0 r}$$

Since there are $\frac{Z(Z-1)}{2}$ pairs of protons, the coulomb energy of the nucleus is given by

$$\begin{aligned} E_c &= \frac{Z(Z-1)}{2} V \\ &= -\frac{Z(Z-1)e^2}{8\pi\epsilon_0} \left(\frac{1}{r}\right)_{av} \end{aligned} \quad (3.59)$$

where $\left(\frac{1}{r}\right)_{av}$ is average value of $\frac{1}{r}$ for all proton pairs. If all protons are uniformly distributed

throughout a nucleus of radius R , $\left(\frac{1}{r}\right)_{av}$ is proportional to $\frac{1}{R}$ and hence to $\frac{1}{A^{1/3}}$, so that the Coulomb energy is given by

$$E_c = -a_3 \frac{Z(Z-1)}{A^{1/3}} \quad (3.60)$$

where a_3 is a constant. This energy is negative as it reduces the stability of the nucleus.

Thus, the total binding energy of the nucleus, as far as analogy to liquid drop model is relevant, is given by

$$\begin{aligned} E_b &= E_v + E_s + E_c \\ &= a_1 A - a_2 A^{2/3} - a_3 \frac{Z(Z-1)}{A^{1/3}} \end{aligned} \quad (3.61)$$

The binding energy per nucleon is, therefore, given by

$$\frac{E_b}{A} = a_1 - \frac{a_2}{A^{1/3}} - a_3 \frac{Z(Z-1)}{A^{4/3}} \quad (3.62)$$

The variation of each term in the above expression with atomic mass number A is shown in Fig. 3.13, along with their sum $\frac{E_b}{A}$. The coefficients were chosen to make the $\frac{E_b}{A}$ curve resemble as closely possible to the empirical binding energy per nucleon curve. The fact that the theoretical curve can be made to agree so well with the empirical one implies that the analogy between a nucleus and a liquid drop has at least some validity. However, this model has limitation and can not explain some of the finer details of the nuclear structure.

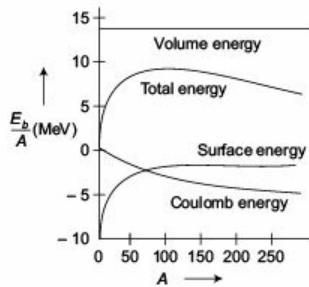


Fig. 3.13 Variation of volume, surface, coulomb and total energies with mass number.

Corrections in the Formula. The binding energy formula given in Eq. (3.61) can be improved by taking into account two effects that do not fit into the simple liquid drop model but which make sense in terms of a model that accounts for the nuclear energy levels. We discuss below these two effects.

Asymmetry Energy. There is a tendency in stable nuclei to have nearly equal numbers of protons and neutrons, especially when A is small. This tendency is often called the *symmetry effect*. The

symmetry effect is a consequence of the Pauli's exclusion principle and the charge independence of nuclear forces. Among any set of isobars (nuclei with same number of nucleons) the nucleus (or nuclei) with Z (number of protons) closest to N (number of neutrons) will have the lowest total energy and hence the most stable.

Suppose that the uppermost neutron and proton energy levels, which the exclusion principle limits to two particles each, have the same energy separation ϵ (Fig. 3.14). In order to produce a neutron excess of, say, $N - Z = 8$, keeping the number of nucleons on the same, $\frac{1}{2}(N - Z) = 4$ neutrons would have to replace protons in an original nucleus with $N = Z$. The new neutrons would occupy levels higher in energy by $2 \epsilon = \frac{4\epsilon}{2}$ than those of the protons they replace. In general, for $\frac{1}{2}(N - Z)$ new neutrons, each must be raised in energy by $\frac{1}{2}(N - Z)\frac{\epsilon}{2}$.

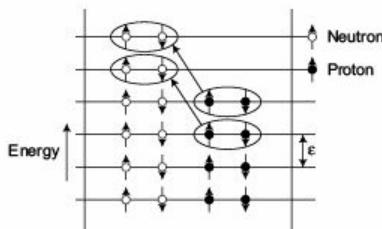


Fig. 3.14 Replacing proton by neutron results in increase in energy level.

The total work needed for this to be done is

$$\Delta E = (\text{Number of new neutrons}) \cdot (\text{increase in energy per new neutron})$$

$$\begin{aligned} &= \frac{1}{2}(N - Z) \cdot \left(\frac{1}{2}(N - Z) \cdot \frac{\epsilon}{2} \right) \\ &= \frac{\epsilon}{8}(N - Z)^2 \end{aligned}$$

As

$$N = A - Z, (N - Z)^2 = (A - 2Z)^2$$

$$\therefore \Delta E = \frac{\epsilon}{8}(A - 2Z)^2 \quad (3.63)$$

The greater the number of nucleons in a nucleus, the smaller is the energy level spacing ϵ , with ϵ proportional to $\frac{1}{A}$. This means that the *asymmetry energy* E_a due to the difference between N and Z can be expressed as

$$E_a = -\Delta E = -a_4 \left(\frac{(A - 2Z)^2}{A} \right) \quad (3.64)$$

where a_4 is a constant. The asymmetry energy is –ve because it reduces the binding energy of the nucleus.

Pairing Energy. The last correction term arises from the tendency of protons and neutrons to exist in pairs. The majority of the stable nuclei (148 to be precise) have both Z and N even; there are 100 nuclei with Z even but N odd or vice versa and there are only 4 nuclei in which both N and Z are odd. This preference for Z and N to be even is explained by a property of the nuclear force called the *pairing effect*. The nuclei such as ^4_2He , $^{12}_6\text{C}$ and $^{16}_8\text{O}$ appear as peaks on the binding energy per nucleon vs A curve. While the odd-odd nuclei with both unpaired protons and neutrons have relatively low binding energies.

The pairing energy E_p is +ve for even-even (both Z and N even) nuclei, O for odd-even and even-odd nuclei and –ve for odd-odd nuclei and is estimated to vary with A as $A^{-4/3}$. Hence,

$$E_p = (\pm, 0) \frac{a_5}{A^{4/3}} \quad (3.65)$$

where a_5 is a constant.

Putting all the five terms together, we obtain *semiempirical binding energy formula* which was first obtained by Weizsacker and is given by

$$E_b = a_1 A - a_2 A^{2/3} - a_3 \frac{Z(Z-1)}{A^{1/3}} - a_4 \frac{(A-2Z)^2}{A} (\pm, 0) \frac{a_5}{A^{4/3}} \quad (3.66)$$

The five coefficients are generally adjusted to fit all the measured binding energies as well as possible. One set of values that gives a good fit is as follows:

$$\begin{aligned} a_1 &= 14.1 \text{ MeV}, & a_2 &= 13.0 \text{ MeV}, & a_3 &= 0.595 \text{ MeV} \\ a_4 &= 19.0 \text{ MeV}, & a_5 &= 33.5 \text{ MeV} \end{aligned}$$

The value of binding energy E_b obtained with Eq. (3.64) agrees better with the observed value than that obtained using Eq. (3.61), which suggests that the liquid-drop model, though a good approximation, does not explain all aspects of the nuclear behaviour.

3.14.2 The Shell Model

The semiempirical binding energy formula obtained in the previous section gives a fairly accurate picture of the general trends of nuclear binding energies. However, the liquid drop model does not show any fluctuation in any property of the nucleus which is based on the energy levels of the nucleons similar to atomic electrons. We know that atomic binding energies fluctuates periodically with atomic number Z , being highest when the electrons completely fill a closed shell and falling drastically just beyond each closed shell.

Magic Numbers. The total numbers of electrons in the closed-shell atoms of noble gas atoms He, Ne, Ar, Kr, Xe, Rn are known to be 2, 10, 18, 36, 54 and 86 electrons respectively, leading to their exceptional stability and chemical inactivity. Therefore, these numbers are sometimes referred to as *atomic magic numbers*.

By the late 1940s, substantial evidence had accumulated that certain nuclei having analogous closed-shell numbers possess unusual stability. These numbers called the *nuclear magic numbers*, are 2, 8, 20, 28, 50, 82 and 126. A nucleus is especially stable if its Z or N is equal to one of these

magic numbers and even more stable if both Z and N are equal to the magic numbers. The magic numbers are the same for protons and neutrons, due to the charge independence of nuclear forces. It is important to note that the nuclear magic numbers are different from the atomic magic number indicating the difference between the dominant forces in nuclei (nuclear forces) and atoms (electrostatic forces).

There are many evidences of the existence of nuclear magic numbers. Nuclei in which Z is magic number, tend to have large numbers of stable isotopes than other nuclei. For example, tin with Z equal to magic number 50 has 10 stable isotopes more than any other element. Similarly, nuclei in which N is a magic number tends to have more stable isotones than average, e.g., $N = 82$

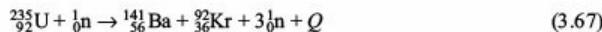
has 7 stable isotones. The binding energy per nucleon $\left(\frac{E_b}{A}\right)$ in each of the 'doubly magic' nuclei $^{20}_{\Lambda}\text{He}_2$, $^{16}_{\Lambda}\text{O}_8$, $^{40}_{\Lambda}\text{Ca}_{20}$, $^{48}_{\Lambda}\text{Ca}_{28}$ and $^{208}_{\Lambda}\text{Pb}_{126}$ is markedly higher than the estimated value using the binding energy formula.

The other evidence of the existence of the magic number is the extremely low neutron capture cross-section of the nuclei with Z or N equal to the magic numbers. Also, in most of the known spontaneous neutron emitters like $^{17}_{\Lambda}\text{O}$, $^{87}_{\Lambda}\text{Kr}$ and $^{137}_{\Lambda}\text{Xe}$ the number of neutron N equals a magic number plus one. This implies small affinity for the extra neutron. This provides yet another evidence for the existence of magic numbers. The existence of nuclear magic number established the existence of the nuclear shell structure.

Theoretical explanation for the nuclear energy levels is possible if the form of the nuclear potential is known. It has been established that the nuclear forces are strongly attractive and short range.

3.15 NUCLEAR FISSION

Nuclear fission is the phenomenon of splitting of a heavy nucleus (usually of mass number greater than 230) into lighter nuclei. For example, nuclear fission of $^{235}_{\Lambda}\text{U}$, when hit by a neutron is represented as



where Q is the energy released in the process. It is found that the sum of the masses of the products is little less than that of the reactants, i.e., a mass defect Δm is created in the process. This mass defect appears in the form of energy released in accordance with Einstein's mass energy relation, $E = \Delta m \cdot c^2$. Most of the energy released appear in the form of kinetic energy of fragments, released neutrons and the γ -rays.

About 188 MeV of energy is released per fission of $^{235}_{\Lambda}\text{U}$ nucleus. This comes out to be 0.8 MeV per nucleon. This is enormous amount of energy and nowadays being exploited for commercial power production.

3.15.1 Chain Reaction

We have seen above that the fission of each $^{235}_{\Lambda}\text{U}$ nucleus produces 3 neutrons. The three neutrons so produced can cause the fission of three more $^{235}_{\Lambda}\text{U}$ nuclei and produce 9 neutrons. These nine neutrons, in turn, can cause the fission of 9 more $^{235}_{\Lambda}\text{U}$ nuclei and produce 27 neutrons. The process keeps on multiplying. Thus, a continuous reaction called *nuclear chain reaction* starts and a huge amount of energy is released in a short interval of time (a few microseconds). Thus, *a chain reaction is the fission process in which the number of neutrons keeps on multiplying rapidly (in*

geometrical progression) during fission till whole of the fissionable material is disintegrated. A chain reaction is depicted in Fig. 3.15.

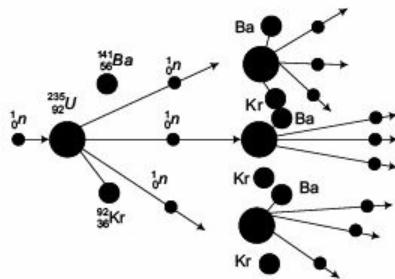


Fig. 3.15

In an atom bomb, an uncontrolled chain reaction proceeds while in a nuclear reactor a controlled chain reaction is allowed to proceed to generate electrical power.

Because in naturally occurring uranium the fraction of fissionable ^{235}U is very less, a chain reaction cannot proceed.

In order to have a sustained chain reaction, it is essential that enough ^{235}U nuclei are available to the neutrons so that the fission process continues. This is possible if fissionable uranium matter is greater than a certain minimum size called the *critical size* and the corresponding mass is called *critical mass*.

All the neutrons produced may not cause further fission. The neutrons may be lost due to: (i) leakage from the system (or fissionable matter) and (ii) capture by impurities or non-fissionable uranium atoms. Whether chain reaction once started will remain steady, increase or decrease depend on a parameter called the *neutron reproduction factor (K)*. This is defined as the ratio of the rate of production of neutrons to the rate of loss of neutrons due to absorption and leakage. That is,

$$\text{Neutron reproduction factor, } K = \frac{\text{Rate of production of neutrons}}{\text{Rate of loss of neutrons}}$$

If $K = 1$, the chain reaction will proceed steadily, i.e., it will be sustained. The size of the fissionable material corresponding to $K = 1$ is called the *critical size*.

If $K > 1$, the rate of chain reaction keeps on increasing resulting in an explosion. This size of the fissionable material is greater than critical size called *super critical size*. This is the working principle of atomic bombs.

Thermal neutrons. For a chain reaction to be sustained, it is required to slow down the neutrons produced in the fission process so that it does not go out of control. For this purpose, secondary fast neutrons are made to pass through substances like heavy water, paraffin or graphite which are rich in hydrogen so that the neutrons slow down. These substances are called *moderators*. On colliding with hydrogen nuclei or protons in the moderator, the neutrons slow down. The slowed down neutrons are referred to *thermal neutrons*. The average energy of a neutron produced in

fission of ^{235}U is 2 MeV, which is reduced to approximately $\frac{1}{40}$ eV due to collision with moderator

atoms. The velocity of thermal neutrons is about 2.2 km/s, which is nearly the same as that of atoms and molecules in a gas at room temperature.

3.15.2 Isotopes of Uranium: Enrichment

Naturally occurring uranium contains the following three isotopes.

Isotope	Abundance	$T_{1/2}$
^{238}U	99.28%	4.51×10^9 years
^{235}U	0.714%	7.1×10^8 years
^{234}U	0.006%	2.48×10^5 years

of these isotopes only ^{235}U can be made to undergo fission even with *slow neutrons* (also called *thermal neutrons*). Two other nuclides which do not occur in nature but have been found to be easily fissionable by neutrons of all energies are ^{233}U and ^{239}Pu . The most abundant isotope ^{238}U requires fast neutrons with energies greater than 1 MeV to undergo fission. Some other nuclei like ^{232}Th and ^{231}Pa also undergo fission only when bombarded by fast neutrons. It has been found that many nuclei can be made to undergo fission if bombarded by highly energetic α -particles, deuterons, neutrons, etc. Bismuth can be made to undergo fission with α -particle of 50 MeV whereas tantalum requires to be hit by α -particles with energy 400 MeV or higher.

For large-scale applications (such as nuclear power plants and nuclear bombs) of fissile materials, ^{233}U , ^{235}U and ^{239}Pu have been found to be useful. Before uranium can be used as a fissile material such as nuclear fuel, the concentration of easily fissionable ^{235}U isotope needs to be increased. This process is often termed as *uranium enrichment*.

3.15.3 Fission Yield Curve: Mass Distribution of Fission Products

In the uranium fission process, barium and krypton are not the only fission products. In fact, a large number of nuclei of intermediate charge and mass are formed. Investigations of the fission products of ^{235}U have shown that the range of their mass numbers vary from 72 to 158. About 97% of the ^{235}U undergoing fission give fragments which fall into two broad groups as shown in the *fission yields curve* in Fig. 3.16. The two groups are:

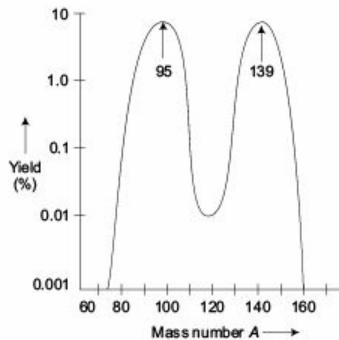


Fig. 3.16 Mass distribution of fragments from fission of ^{235}U .

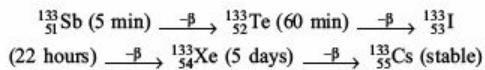
(i) *Light group* with mass numbers from 85 to 104, and

(ii) *Heavy group* with mass numbers from 130 to 149.

The most probable type of fission which occurs in about 7% of the total cases, gives fission products with mass numbers 95 and 139.

3.15.4 Fission Decay Series

Normally fission fragments have too many neutrons in their nuclei to be stable. Consequently, most of them decay by electron emission. Each fragment starts a short radioactive series involving many emissions of β -particles. These series are called *fission-decay series* and each series has three members (decays) on the average, although longer and shorter chains occur frequently. One such fission decay chain is shown below which starts with one of the unstable fragments of the fission of ^{235}U nucleus:



In the bracket the mean life time of the nucleus is mentioned.

3.15.5 Neutron Emission in Fission

One of the notable features of nuclear fission is that while it is initiated by neutrons, it is also accompanied by the emission of fast moving neutrons, with velocity as high as $1.6 \times 10^7 \text{ m/s}$. The number of neutrons released depends on the mode of fission and the energy of the neutrons which induce fission. The average number of neutrons emitted per thermal neutron absorbed by the three important fissile materials are: $^{235}\text{U} \rightarrow 2.43$, $^{233}\text{U} \rightarrow 2.50$ and $^{239}\text{Pu} \rightarrow 2.89$. These neutrons are emitted by the fission fragments and not by the compound nucleus. The neutrons emitted as a result of fission process can be divided into the following two groups:

- (i) **Prompt Neutrons.** These neutrons are ejected by the product nuclei within 10^{-14} second of the fission process. These neutrons are about 99.36% of the total fission neutrons.
- (ii) **Delayed Neutrons.** These neutrons are about 0.64% of the total neutrons due to the fission of ^{235}U . These are emitted with gradually decreasing intensity for several minutes after the actual fission process. Although, the number of delayed neutrons is small, they have a strong influence on the time dependent behaviour on the chain reaction systems based on fission and play an important role in the control of nuclear reactors based on the fission process.

3.16 NUCLEAR REACTOR

In a nuclear reactor a controlled self-sustaining chain reaction of nuclear fission process takes place producing a large amount of heat energy. Nuclear reactors are used for the following purposes:

- (i) To generate electrical power.
- (ii) To produce radioactive isotopes which are used in the fields of medicine, agriculture, industry, etc.
- (iii) To propel ships, submarines, etc.

- (iv) To produce bomb grade fissile materials like ^{239}Pu .
- (v) Research, e.g., to produce neutron beam of high intensity.

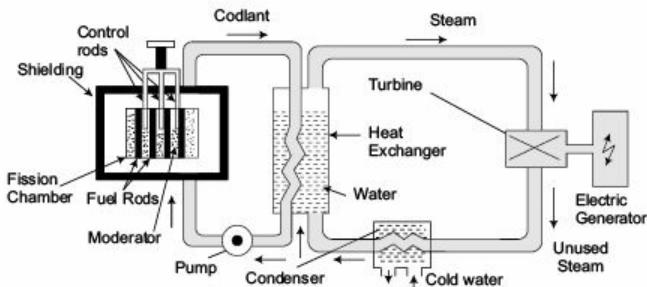


Fig. 3.17 Schematic diagram of a nuclear reactor.

Construction. A labelled diagram of a nuclear reactor is shown in Fig. 3.17. The main components of a nuclear reactor are:

- (i) **Nuclear fuel.** Generally enriched ^{235}U (containing higher percentage of ^{235}U than natural uranium), sealed in aluminium (or stainless steel or zirconium) cylinders is used as fuel. (Sometimes ^{233}U or ^{239}Pu is also used as fuel). These cylinders are arranged in fission chamber inside the reactor. Space is provided between individual rods to allow free passage of the coolant. When slow neutrons bombard on the fuel, fission starts. Periodically the rods are replaced as they get contaminated by the fission products.
- (ii) **Moderator.** The neutrons released by the fission or uranium have high energies of the order of 2 MeV. However, for a controlled chain reaction, slowly moving neutrons are required. Moderators are the substances (rich in protons) used to slow down the fast moving neutrons. Heavy water (D_2O), graphite, beryllium (or beryllium oxide), paraffin, water, etc. are some of the commonly used moderators.

When fast moving neutrons collide with the protons of the moderator, the neutrons are slowed down. Such neutrons are called thermal neutrons which on striking cause fission of ^{235}U nucleus in the fuel rod.

The Apsara reactor at BARC uses water as moderator, while reactors at Tarapur, Kalpakkam, Kota and Narora use heavy water as moderator.

- (iii) **Control rods.** To control the chain reaction, rods of neutron absorbing materials like boron or cadmium are inserted into the reactor core. By suitably adjusting the height of the control rods inside the chamber, only desired number of neutrons are left to produce fission. Whenever, the reactor is to be shut down, the control rods are inserted to a considerable depth so that K becomes less than unity and the chain-reaction can no longer be maintained. Sometimes in addition, safety rods are provided which are used to stop the nuclear fission by inserting it into the reactor, i.e., to shut down the reactor.
- (iv) **Coolant.** A large amount of heat is produced due to nuclear fission inside the reactor. Coolant is used to remove the intense heat produced and to bring it out for being utilized.

Water and heavy water are good coolants. However, due to their low boiling points, they need pressurisation. At high temperature liquid metals like molten sodium or potassium are excellent coolants. Gaseous (CO_2) coolant is used in large-power reactors.

The coolant releases the heat energy to the water in a heat exchanger producing superheated steam which drives the turbines of generator.

- (v) **Shielding.** The whole reactor is enclosed in a thick (2 to 3 m) concrete wall called protective shield. It prevents thermal and nuclear (γ and neutron) radiations from coming out and protect the people in immediate vicinity.

Working. In the beginning, some neutrons are produced by the action of α -particles on beryllium. These neutrons are slowed down and are used to initiate fission of ^{235}U nuclei. Fast neutrons given out in the fission process are slowed down by the moderator. These thermal neutrons are then used to carry out sustained-chain reaction with the help of control rods. Heat generated in the process is carried out by the coolant for heating water in the heat exchanger. The steam so produced is used to drive the turbines of a generator.

3.17 NUCLEAR FUSION

Nuclear fusion is the process in which two or more lighter nuclei fuse to form a single stable and heavy nucleus. For example, when two deuterium nuclei fuse together, a helium nucleus is formed. This is represented as



where Q is the energy released in the process. The mass of the product nucleus is slightly less than the masses of the lighter nuclei fusing together. This difference in masses (called mass defect Δm) results in the release of tremendous amount of energy, in accordance with Einstein's mass energy relation, $E = \Delta m \cdot c^2$.

Condition for nuclear fusion. The fusion reactions take place under the conditions of extremely high temperature ($\approx 10^7$ K) and pressure. This is necessary so that protons have high enough kinetic energy to overcome their mutual repulsion and to come closer than the range of nuclear force. Also, a higher density is needed so that collisions between the lighter nuclei are frequent. Such conditions exist in the interior of the sun and other stars and it is believed that the energy liberated in the sun and other stars is due to the nuclear fusion reactions.

Because nuclear fusion reaction occurs at extremely high temperature, it is also known as *thermonuclear reaction*.

Due to the difficulties in creating the necessary conditions in a laboratory, it has not been possible to carry out controlled nuclear fusion. However, hydrogen bomb, based on the fusion of hydrogen atoms into heavier atoms, have been tested by many countries. In hydrogen bombs, first an atom bomb (based on fission) is exploded to create the necessary condition of high temperature and pressure. If controlled nuclear fusion can be carried out, all the energy related problems in the world can be solved because unlike fission based nuclear power plants which produces harmful nuclear waste and radiations, clean unlimited nuclear power can be produced.

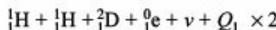
Stellar energy. Energy in stars including the sun is produced by nuclear fusion. The necessary conditions for the nuclear fusion to take place are available inside stars. Based on several studies, it has been found that hydrogen is the most abundant material inside stars. H.A. Bethe proposed that

inside stars the fusion of hydrogen nuclei to form helium takes place. In this process, mass defect Δm appears in the form of energy in accordance with mass-energy relation, $E = \Delta m \cdot c^2$.

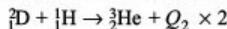
Two types of thermonuclear reactions have been proposed in the stars for fusion of four hydrogen nuclei to form a helium atom along with the release of huge amount of energy. These are: (i) Proton-Proton ($p - p$) cycle and (ii) Carbon-Nitrogen ($C - N$) cycle.

It is believed that in the initial phase of the life of a star, the ($p - p$) cycle is dominant. Later when the temperature of the central core becomes higher and concentration of helium increases, as in the present state of the sun, the $C - N$ cycle is the main source of energy.

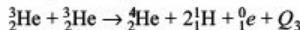
Proton-Proton cycle. It is so called because the first step involves the combination of two protons (or hydrogen nuclei). When two protons fuse, they produce a deuteron.



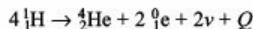
The deuteron then combines with another proton to yield helium-3.



The two helium-3 nuclei fuse together to produce helium-4.

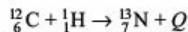


For this reaction to occur, clearly, the first two reactions occur twice. The above reactions can be summed to obtain

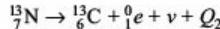


The value of Q , as calculated from the mass defect, is 26.7 MeV. Thus, the net reaction is that four hydrogen atoms fuse together to produce a helium atom and 26.7 MeV of energy is released.

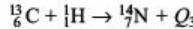
Carbon-Nitrogen cycle. In this cycle, carbon acts as a nuclear catalyst. It consists of the following reactions. To begin with, a proton fuse with carbon to form ${}^{13}_7N$



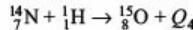
The product ${}^{13}_7N$ is radioactive with half-life of 10 minutes. It decays into ${}^{13}_6C$ as



The stable ${}^{13}_6C$ nuclei fuse with another proton as



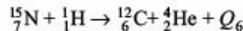
The stable ${}^{14}_7N$ combines with another proton



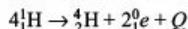
The unstable ${}^{15}_8O$ (half-life 2.06 minutes) decays according to the process



Finally, ${}^{15}_7N$ reacts with the fourth proton, as



The above reactions can be summed to obtain



In this case also, Q is calculated to be 26.7 MeV.

3.18 NUCLEAR HOLOCAUST

If used for peaceful purposes, nuclear source of energy is an important source because thermal source of energy is exhaustible as after some decade later there will be no coal and oil to produce thermal power. However, if used as weapon in nuclear warfare, it has potential of causing tremendous destruction. This potential threat is often referred to as *nuclear holocaust*.

We have seen that a single uranium fission releases about 190 MeV of energy instantly. If each nucleus of about 50 kg of $^{235}_{92}\text{U}$ undergoes fission, about 4×10^{15} joule of energy would be released. This is equivalent to about 20,000 tons of TNT, enough for a super explosion. Uncontrolled release of large amount of energy is said to be an *atomic explosion*. The first atomic explosion in the world occurred on August 6, 1945 when USA dropped an atom bomb on Hiroshima, Japan, causing large-scale devastation. The explosion was equivalent to 20,000 tons of TNT. The explosion killed 66,000 people out of 3,43,000 persons in the city spread over 10 sq. km while 69,000 people were injured. About 67% of the city's structure were destroyed.

Hydrogen bombs, based on fusion of hydrogen, are more powerful. One such bomb was tested in 1954 was equivalent to 10 megatons of TNT. Since then many countries, including India, have tested this type of bomb.

It is estimated that a nuclear arsenal sufficient to destroy the earth several times exists in the world. A nuclear holocaust will not only destroy all forms of life on the earth, but its radioactive fallout will make the earth unfit for life for all time to come.

It is predicted that in the event of a nuclear holocaust, the radioactive waste will hang like a cloud in the earth's atmosphere. It will absorb the sun's radiation resulting in, what is called, *nuclear winter*.

Let us pray that wisdom prevail upon the leaders of the world so that nuclear energy is used in improving the quality of life on the earth and not for destroying it.

Radiation Hazards. Nuclear radiations are very harmful for human beings. When exposed to nuclear radiations, the complex organic molecules of the human organisms break up and therefore the normal functioning of the biological system is disrupted. This leads to permanent damage of the tissues and ultimately even to death.

The extent of damage due to nuclear radiations depends on the following factors: (i) dose of radiation, (ii) rate of dose and (iii) the nature of organism exposed.

The unit used to measure the dose (or quantity) of radiation is *roentgen (R)*. A single dose of one roentgen is defined to be the radiation which produces 1.6×10^{12} pairs of ions in 1 gm of air or 2.08×10^9 pairs of ions in 1 c.c of air. One thousandth part of the dose is a milliroentgen (mR). Obviously, this unit called *exposure* (or dosage) unit is independent of time. The radiation absorbed per unit time is called *exposure rate* or *dosage rate* or *radiation absorbed dose*.

The *permissible dose* or *safe limit* of exposure to radiation over the entire body is about 250 mR per week. The over exposure causes the damage to human organism. The damage caused due to the exposure to nuclear radiations, can be of two types:

- (i) Pathological and (ii) Genetic.

In the *pathological damage*, the over exposed ($> 600 R$) organism may ultimately die. Smaller dose of about 100 R may cause start of leukemia (death of Red Blood Corpuscles in the blood) or cancer.

In the *genetic damage*, the radiations cause injury to genes in the reproductive cells. This causes mutations of the chromosomes of the cellular nuclei. Mutations are harmful, irreversible and pass on from generation to generation. Several families in Hiroshima, Japan, are still suffering from genetic damage even after three generations.

Let us pray and hope that the nuclear energy is used for the welfare of the human beings and not to destroy them.

Example 3.14

How many fissions take place per second in a 300 MW reactor? Assume that each fission releases 200 MeV of energy.

Solution. Let n be the number of fissions per second.

$$\text{Energy released per second} = 200 \times 10^6 \text{ eV} \times 1.6 \times 10^{-19} \frac{\text{J}}{\text{eV}} \times n$$

$$= n \times 3.2 \times 10^{-11} \text{ J}$$

$$\text{Power} = \frac{\text{Energy}}{\text{Time}}$$

$$300 \times 10^6 \text{ Js}^{-1} = n \times 3.2 \times 10^{-11} \text{ J}$$

$$\text{or } n = \frac{300 \times 10^6}{3.2 \times 10^{-11}} = 9.38 \times 10^{18}$$

Example 3.15

The fission of a single nucleus of ^{235}U releases 200 MeV of energy. Calculate the number of fission that must occur to produce a power of 1 MW.

Solution. Let the number of fissions per second be n .

$$\begin{aligned} \text{Energy released per second} &= n \times 200 \text{ MeV} \\ &= n \times 200 \times 1.6 \times 10^{-13} \text{ J} \end{aligned}$$

$$\text{Energy to be produced per second} = \text{Power} \times \text{Time}$$

$$\begin{aligned} &= 1 \text{ MW} \times 1 \text{ s} \\ &= 10^6 \text{ J} \end{aligned}$$

$$\therefore n \times 200 \times 1.6 \times 10^{-13} \text{ J} = 10^6 \text{ J}$$

$$\text{or } n = \frac{10^6}{200 \times 1.6 \times 10^{-13}} = 3.125 \times 10^{16}.$$

Example 3.16

Calculate the energy released by the fission of 5 gm of ^{235}U in kWh if energy released per fission of ^{235}U is 200 MeV.

Solution. Number of atoms in 5 gm of ^{235}U = $\frac{\text{Avogadro's number}}{\text{Mass number}} \times 5$

$$= \frac{6.023 \times 10^{23}}{235} \times 5$$

Energy released by the fission of 5 gm of ^{235}U

$$\begin{aligned}
&= \frac{6.023 \times 10^{23}}{235} \times 5 \times 200 \text{ MeV} \\
&= \frac{6.023 \times 10^{23}}{235} \times 5 \times 200 \times 1.6 \times 10^{-13} \text{ J} \\
&= 4.1 \times 10^{11} \text{ J} \\
&= \frac{4.1 \times 10^{11}}{3.6 \times 10^6} \text{ kWh} \\
&= 1.139 \times 10^5 \text{ kWh.}
\end{aligned}$$

Example 3.17

The neutron separation energy is defined as the energy required to remove a neutron from a nucleus. Calculate the neutron separation energy of the nucleus $_{20}^{41}\text{Ca}$ from the following data:

$$\begin{aligned}
m ({}_{20}^{40}\text{Ca}) &= 39.962591 \text{ u}, \\
m ({}_{20}^{41}\text{Ca}) &= 40.962278 \text{ u} \text{ and } m ({}_{0}^{1}\text{n}) = 1.008665 \text{ u}
\end{aligned}$$

Solution. The neutron separation from ${}_{20}^{40}\text{Ca}$ can be represented as



Mass defect in this process,

$$\begin{aligned}
\Delta m &= m ({}_{20}^{40}\text{Ca}) + m ({}_{0}^{1}\text{n}) - m ({}_{20}^{41}\text{Ca}) \\
&= 39.962591 \text{ u} + 1.008665 \text{ u} - 40.962278 \text{ u} \\
&= 40.971256 \text{ u} - 40.962278 \text{ u} \\
&= 0.008978 \text{ u}
\end{aligned}$$

\therefore Neutron separation energy,

$$\begin{aligned}
Q &= 0.008978 \times 931 \text{ MeV} \\
&= 8.3585 \text{ MeV.}
\end{aligned}$$

Example 3.18

It is estimated that the atomic bomb exploded at Hiroshima released a total energy of $7.6 \times 10^{13} \text{ J}$. If on the average 200 MeV energy was released by the fission of one ${}^{235}\text{U}$ atom, calculate (i) the number of uranium atoms fissioned, and (ii) the mass of uranium used in the bomb.

Solution. (i) Number of ${}^{235}\text{U}$ atoms fissioned,

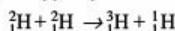
$$\begin{aligned}
n &= \frac{\text{Total energy released}}{\text{Energy released per fission}} \\
&= \frac{7.6 \times 10^{13} \text{ J}}{200 \times 1.6 \times 10^{-13} \text{ J}} = 2.375 \times 10^{24}
\end{aligned}$$

(ii) Mass of uranium used,

$$\begin{aligned}
m &= \frac{\text{Mass Number}}{\text{Avogadro's number}} \times n. \\
&= \frac{235 \times 2.375 \times 10^{24}}{6.023 \times 10^{23}} = 926.65 \text{ gm}
\end{aligned}$$

Example 3.19

Calculate the energy released per gram of fuel for the reaction



Given: $m({}_1^2\text{H}) = 2.014102 \text{ u}$, $m({}_1^3\text{H}) = 3.01609 \text{ u}$, $m({}_1^1\text{H}) = 1.007825 \text{ u}$

$$\begin{aligned}\text{Solution. Energy released per fusion} &= (2 \times 2.014102 \text{ u} - 3.01609 \text{ u} - 1.007825 \text{ u}) \times 931.5 \frac{\text{MeV}}{\text{u}} \\ &= 4.033 \text{ MeV}\end{aligned}$$

$$1 \text{ gm of fuel has } \frac{0.001 \text{ kg}}{2 \times 1.66 \times 10^{-27} \text{ kg}} = 3 \times 10^{23} \text{ nuclei}$$

Two deuterium nuclei release 4.033 MeV of energy.

$$\begin{aligned}\therefore \text{Energy released per gram} &= 3 \times 10^{23} \times \frac{4.033}{2} \\ &= 6.05 \text{ MeV}\end{aligned}$$

Example 3.20

Estimate the temperature required for the deuterium-tritium fusion reaction to occur if the kinetic energy of each particle is 0.225 MeV. The radius of deuterium nucleus $\approx 1.5 \text{ fm}$ and that of tritium $\approx 1.7 \text{ fm}$.

Solution. Assume that the two nuclei collide head on, the separation between the two nuclei is equal to the sum of their radii. The electrostatic potential energy of the two particles at this distance must be equal to the kinetic energy of the two when they are far apart.

$$\begin{aligned}2 \text{ KE} &= \frac{ke^2}{r_d + r_t} = \frac{8.99 \times 10^9 \text{ Nm}^2 \text{ C}^{-2} \times (1.6 \times 10^{-19} \text{ C})^2}{3.2 \times 10^{-15} \text{ m} \times 1.6 \times 10^{-19} \text{ J/eV}} \\ &= 0.45 \text{ MeV}\end{aligned}$$

Kinetic energy of each particle $\approx 0.225 \text{ MeV}$

$$\text{Average kinetic energy} = \frac{3}{2} kT \quad (k = \text{Boltzmann Const.})$$

$$\begin{aligned}\therefore T &= \frac{2 \times \text{average kinetic energy}}{3k} \\ &= \frac{2 \times 0.225 \text{ MeV} \times 1.6 \times 10^{-13} \text{ J/MeV}}{3 \times 1.38 \times 10^{-23} \text{ J/K}} \\ &= 2 \times 10^9 \text{ K}\end{aligned}$$

3.19 ACCELERATORS

The most important source of information about nuclei is the *scattering* or *collision experiments*, in which a beam of energetic projectiles, such as α -particle or protons, neutrons, etc. is directed at a target containing the nuclei to be studied. The projectiles may be scattered by the target or induce

various reactions, but in either case one may conclude about the properties of the target. Many important discoveries such as the discoveries of the nucleus, neutron, fusion, etc. were possible due to the collision experiments. For this purpose, the projectiles are needed to be *accelerated* to high energies. We discuss below the *accelerator – the machines used to accelerate the projectiles to high energy so as to cause meaningful collisions*.

The earliest collision experiments used α -particles produced by naturally occurring radioactive nuclei. These α -particles possessed energy in the range of 4 MeV to 9 MeV. By 1920s, the need of highly energetic alpha particles were felt. This prompted several physicists to design accelerators. Since then many different accelerators have been developed, some of which we discuss below. All these accelerators use electric field to accelerate protons, electrons and the positive and negative ions.

3.19.1 DC Accelerators

The most easiest way of accelerating a charged particle is to use an electric field applied along the length of an evacuated tube, as shown in Fig. 3.18. The field is produced by applying a potential difference between the ends of the tube. Because the polarity is not reversible, the accelerator is a *dc accelerator*.

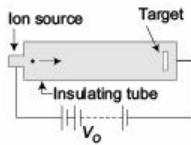


Fig. 3.18 A dc accelerator with applied voltage V_0

To find out anything about the nuclear structure, the projectile must approach the target nucleus within 1 to 2 fm, closer the better. Because of the Coulomb repulsion (assuming the projectile is positively charged) between the projectile and the nucleus, the projectiles should be highly energetic, e.g., to bring two protons within 1 fm, requires the projectile proton to possess 1.5 MeV of energy. Therefore, the accelerator should give the projectile a kinetic energy in MeV. If a projectile with charge q is accelerated through a potential difference V_0 , the kinetic energy of the projectile is $K = q V_0$. To produce energy in MeV, V_0 should be in megavolt. In 1920s, there was no known way to produce such large dc voltage.

The first successful high-voltage accelerator was produced by English physicists Cockcroft and Walton, who used ac transformers and rectifiers to apply a dc voltage across a bank of capacitors. In this way they accelerated protons to 500 keV and in 1932 could produce the first artificial nuclear reaction. The modern versions of the Cockcroft-Walton accelerator can achieve about one million volts and are widely used to inject projectile into higher energy accelerators such as linear accelerator.

In early 1930s, American physicist Robert van de Graaf successfully designed an electrostatic generator. In this machine, the high voltage terminal is a large hollow conductor, usually spherical or oval in shape. A motor driven belt carries electrostatic charges to the inside of this terminal, where they are deposited and immediately move to the outside surface. In this way a large total charge can

be placed on the terminal and produce a correspondingly high voltage. Modern versions enclose the entire device in a tank of high pressure gas to inhibit loss of charge by sparking. In this way voltage upto 10 MV can be achieved which is used to accelerate many particles, ranging from electrons to heavy ions such as uranium.

3.19.2 AC Accelerators

In a dc accelerator particles are accelerated in a single large boost which requires a very high voltage for any significant acceleration. An alternative to this is to accelerate the projectile particles in steps, in a succession of smaller boosts. In this way a higher final energy can be given to the accelerating particles without the application of a too high voltage. Most of the accelerators in use nowadays are of this multiple boost type. All such accelerators use alternating potentials, and therefore they are referred to as *ac accelerators*. In these accelerators high frequency alternating potential in the range of 10 to 200 MHz (radio frequency range) is used.

3.20 LINEAR ACCELERATOR (LINAC)

A linear accelerator, or linac, consists of a series of metallic cylinders of suitably increasing lengths, arranged coaxially with accelerating gaps between them, placed in a long evacuated chamber (Fig. 3.19). Alternate sets of cylinders are connected together and an alternating voltage supply is connected between the two sets, having a frequency such that the potentials of the tubes oscillate in such a way that whenever the particle reaches a gap between two tubes, there is an accelerating field in the gap. The successive cylinders are made longer so the accelerating particle takes the same time to traverse each cylinder in spite of its increased velocity. Thus, the time taken to pass through each cylinder must be half of the time period of the applied alternating potential and the accelerating particle reach the gap when the next tube is at the maximum potential.

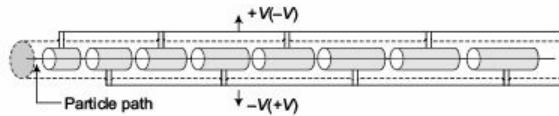


Fig. 3.19 Schematic diagram of a linear accelerator.

If the charge on the accelerating particle is q and the applied potential is V , when the particle moves from one tube to the next it acquires an energy of qV . If there are N gaps in the accelerator, the particle gains an energy NqV in its passage through the whole accelerator. In most linacs the particle is initially accelerated by an alternative arrangement (e.g., a Cockroft-Walton accelerator) so that when it enters the first tube it already possess a kinetic energy K_0 so that the total kinetic energy $K = K_0 + NqV$. If the final velocity acquired by the particle of mass m is v , then

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

or $v = \sqrt{\frac{2K}{m}}$ (3.69)

provided v is small compared to c , the velocity of light. It follows from Eq. (3.69) that for the ions of constant charge and mass, the velocity is proportional to \sqrt{V} .

The underlying principle is that as long as a charged particle is inside the metallic (conducting) tube, it is not influenced by the potential of the tube since the field inside a conductor is zero, but as soon it emerges it is accelerated. Each successive tube is designed in the same way that the accelerating particle take a half-period (of the applied ac potential) to traverse the tube and hence arrive at the next gap at the maximum potential drop in the next tube. A typical frequency for the ac supply is of the order of 100 MHz for which half period is 5×10^{-9} s. In this time an accelerating proton with about 100 MeV energy, $v \approx 0.5 c$ travels nearly 1 m. Therefore, linear accelerators, sometimes with hundreds of tubes can be very long machine.

The design considerations of linear accelerator for electrons are slightly different from proton linear accelerator, as the electrons being lighter attains speed comparable to the speed of light quickly. Therefore, the tubes in electron linac are longer than those in proton linac. The electron linac at Stanford, USA, is over 3 km long and can accelerate electrons upto 50 GeV. Nowadays, many hospitals have 10 MeV linacs to accelerate particles to irradiate tumours.

The main advantage of linac over magnetic accelerating machine is the well collimated high beam intensity of the accelerated particles.

3.21 CYCLOTRON

To obtain high energy particles, a linac must have large number of tubes making the machine too long. In 1929, American physicist Ernest Lawrence thought of using magnetic field to make the accelerating particles move in circles. Lawrence designed an accelerator, called cyclotron, for which he was awarded Nobel Prize in physics in 1939.

The schematic diagram of a cyclotron is shown in Fig. 3.20. It consists of a large evacuated cylindrical box between the pole pieces of an electromagnet. Inside the box, two hollow D-shaped metal cavities, called 'dees' are placed with the open sides slightly apart. The dees are connected to the terminals of an alternating high frequency voltage source which produce an alternating high frequency voltage which produces an alternating electric field in the gap between the dees. When one dee is positive, the other is negative and vice versa. The electric field inside the dee is zero due to the electrical shielding effect. The direction of the magnetic field is perpendicular to the cross-sectional area of the dees.

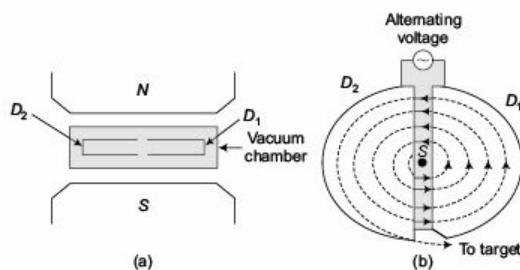


Fig. 3.20 (a) Schematic diagram of a cyclotron and (b) the dees and the circular orbits.

The ions are produced in an electric discharge placed close to the centre of the dee. The gas whose nuclei are to be accelerated (e.g., hydrogen for protons, deuterium for neutrons, helium for alpha particles, etc.) is fed to the electric discharge. Let $+q$ be charge on the particle to be accelerated and m its mass. Suppose D_1 is positive and D_2 negative at the instant the ion is produced from the source S . The ion is accelerated towards D_2 , gaining an energy qV , where V is the potential on D_2 . Once it enters, D_2 is not influenced by the electric field but the magnetic field is present which makes the ion move in a circular path inside D_2 . The frequency of the alternating applied potential is so adjusted that when the ion comes out of D_2 , the polarities of the dees change, i.e., D_1 becomes negative and D_2 positive. The ion is again accelerated towards D_1 , gaining additional energy qV . Again it moves in a circular path inside D_1 . Thus, the ion gains an energy qV each time it enters from one dee to the other. As the velocity of the ion increases, the radius of its circular path also increases. Eventually when the ion approaches the outer edge of the dee, a pair of charged plates deflect the ion out of the machine towards the target.

The force experienced by a particle of charge q moving with velocity v in a magnetic field B is qvB . Since the particle moves in a circle, a centripetal force $\frac{mv^2}{r}$, (r = radius of the circle) must act on it. Thus,

$$qvB = \frac{mv^2}{r} \quad (3.70)$$

$$\text{or} \quad v = \frac{qBr}{m} \quad (3.71)$$

The condition given by Eq. (3.70) is often called the *resonance condition*. The time required to complete a circular revolution is given by

$$T = \frac{2\pi r}{v} = \frac{2\pi m}{qB} \quad (3.72)$$

Hence, the frequency of revolution f is given by

$$f = \frac{1}{T} = \frac{qB}{2\pi m} \quad (3.73)$$

This is also known as *cyclotron frequency*. If the frequency of the ac supply is fixed at the cyclotron frequency, the ions continue to enter and leave in time with the maximum accelerating field throughout the acceleration process. The fact that a cyclotron can be driven by an ac supply of fixed frequency was a major simplification and made it a practical accelerator used in many nuclear laboratories around the world.

The maximum proton energy attainable in a conventional cyclotron is of the order of 20 MeV. At higher energies, Eq. (3.73) is not strictly valid due to the relativistic considerations, i.e., increase in the mass of the particle with increasing velocity has to be considered. As the mass increases, the frequency of the alternating voltage must be reduced to keep the resonance condition. Hence, an ordinary cyclotron cannot accelerate heavy ions to very high energies. For electrons, the relativistic effect becomes more pronounced at much lower energies. Therefore, an ordinary cyclotron cannot be used for accelerating electrons.

Synchrocyclotron or Frequency Modulated Cyclotron

At speeds comparable to the speed of light, m needs to be replaced by the relation $m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$,

where m_0 is the rest mass of the accelerating particle. Therefore, Eq. (3.73) can be expressed as

$$f = \frac{qB}{2\pi m_0} \cdot \sqrt{1 - \frac{v^2}{c^2}} \quad (3.74)$$

Thus, to maintain the resonance (frequency) condition keeping magnetic field B constant, the frequency of the applied field has to be gradually decreased with increasing mass of the ion. A cyclotron, in which the magnetic field B is kept constant and the frequency is decreased continuously is called synchrocyclotron or frequency modulated (FM) cyclotron. In such a cyclotron, a continuous beam is not possible. Hence, the particles are accelerated in short pulses. The frequency is varied from the maximum value when the particles are near the centre to a minimum value when the particles are deflected out of the cyclotron. For a typical synchrocyclotron the frequency range is 36-18 MHz. The energy of the protons is about 750 MeV at a field strength of 2.3 T.

Example 3.21

A cyclotron of maximum radius 0.25 m accelerates protons. Calculate the frequency of the applied ac voltage if the applied magnetic field is 2.0 T. Also find the kinetic energy of the accelerated protons.

Solution. Here, $r = 0.25$ m, $B = 2$ T, $m = 1.67 \times 10^{-27}$ kg

$$\therefore \text{Frequency} \quad f = \frac{qB}{2\pi m} = \frac{1.6 \times 10^{-19} \text{ C} \times 2T}{2 \times 3.14 \times 1.67 \times 10^{-27} \text{ kg}} = 30.5 \text{ MHz}$$

$$\begin{aligned} \text{Kinetic energy} \quad K &= \frac{1}{2} mv^2 = \frac{q^2 B^2 R^2}{2m} \\ &= \frac{(1.6 \times 10^{-19} \text{ C})^2 \times 4 \text{ T}^2 \times (0.25 \text{ m})^2}{2 \times 1.67 \times 10^{-27} \text{ kg}} = 0.19162 \times 10^{-11} \text{ J} \\ &= 12 \text{ MeV} \end{aligned}$$

Example 3.22

A cyclotron accelerates protons to 10 MeV energy (and hence increases v to 4.38×10^7 m/s) with a maximum orbital radius of 40 cm. (a) Calculate the value of the applied magnetic field. (b) Find the frequency of the applied ac supply. (c) If the amplitude of the accelerating voltage is 50 kV, how many complete orbits are made in process of the acceleration?

Solution. (a) Here, $m = 1.67 \times 10^{-27}$ kg, $v = 4.38 \times 10^7$ m/s

$$q = 1.60 \times 10^{-19} \text{ C}, r = 40 \text{ cm} = 0.4 \text{ m}$$

$$\therefore B = \frac{mv}{qr} = \frac{1.67 \times 10^{-27} \times 4.38 \times 10^7}{1.60 \times 10^{-19} \times 0.4} = 1.14 \text{ T}$$

$$(b) \text{ Frequency of ac supply } f = \frac{qB}{2\pi m}$$

$$= \frac{1.60 \times 10^{-19} \times 1.14}{2 \times 3.14 \times 1.67 \times 10^{-27}} = 17.4 \times 10^6 \text{ Hz}$$

$$= 17.4 \text{ MHz}$$

(c) Total energy gained = 10 MeV

Energy gained in one orbit, $qV = 50 \text{ keV}$

$$\therefore \text{Number of complete orbits, } N = \frac{10 \text{ MeV}}{50 \text{ keV}} = 100 \text{ orbits.}$$

Example 3.23

A projectile is suitable to probe a nucleus, if the wavelength of the projectile is of the same order as the size of the nucleus. Is a 30 MeV proton or 30 MeV α -particle better suited for investigating nucleus?

Solution. We know, $E = \frac{p^2}{2m}$ or $p = \sqrt{2mE}$

Wavelength $\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mE}}$

For proton: $\lambda_p = \frac{6.62 \times 10^{-34}}{\sqrt{2 \times 1.67 \times 10^{-27} \times 30 \times 10^6 \times 1.6 \times 10^{-19}}} = 5.2 \times 10^{-15} \text{ m}$

For alpha particle, $\lambda_\alpha = \frac{6.62 \times 10^{-34}}{\sqrt{2 \times 4 \times 1.6 \times 30 \times 10^{-27} \times 30 \times 10^6 \times 1.6 \times 10^{-19}}} = 2.6 \times 10^{-15} \text{ m}$

Size of the nucleus $\approx 2r = 2 \times 1.2 \times 10^{-15} \text{ m} = 2.4 \times 10^{-15} \text{ m}$

Thus, the α -particle is better suited to probe the nucleus as its wavelength is of the order of the size of nucleon.

3.22 BETATRON

We have seen that in the accelerators discussed earlier there is problem in accelerating electrons. An accelerator for accelerating electrons specially was designed by D.W. Kerst in 1941 at University of Illinois. Beams of energetic electrons are needed for various purposes such as initiating nuclear reactions and production of X-rays.

A betatron is an electron accelerator capable for producing electron beams of high energy as well as X-rays of extremely high penetrating power. This ingenious device differs from the cyclotron in at least two fundamental respects. First, the electrons are accelerated by a rapidly changing

magnetic field, and secondly, the circular orbit of the particles has a constant radius.

To understand the working principle of a betatron, consider Fig. 3.21, which represents an electron moving in an orbit of radius r , where the total magnetic flux through the orbit is ϕ and flux density at orbit itself is B , the manetic field lines being perpendicular to the plane of the orbit.

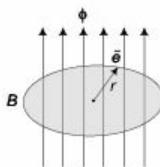


Fig. 3.21

The flux ϕ is increased at a rate $\frac{d\phi}{dt}$. As a consequence, the e.m.f. induced in the orbit is given by

$$\varepsilon = - \frac{d\phi}{dt} \quad (3.75)$$

The work done on the electron in one revolution is, therefore, given by $e\varepsilon$. This must be equal to the tangential force F acting on the electron multiplied by the orbital path $2\pi r$. That is,

$$e\varepsilon = F \cdot 2\pi r \quad (3.76)$$

or $F = \frac{e\varepsilon}{2\pi r} = - \frac{e}{2\pi r} \frac{d\phi}{dt}$ (3.77)

When the electron energy increases because of this applied force, it will attempt to move to an orbit of larger radius. This must be resisted if a stable orbit of constant radius is to be maintained.

The inward radial force on the electron is evB , which must be equal to the centripetal force acting on the electron moving away from the centre. Thus,

$$evB = \frac{mv^2}{r} \quad (3.78)$$

or $p = m v = eBr$ (3.79)

The tangential force F on the electron equals its rate of change of momentum $\frac{dp}{dt}$, which is given by

$$F = \frac{dp}{dt} = er \frac{dB}{dt} \quad (3.80)$$

for an orbit of constant radius.

In order that the radius of the orbiting electron remains constant, the two values of force obtained above must be equal, i.e.,

$$\frac{e}{2\pi r} \cdot \frac{d\phi}{dt} = er \frac{dB}{dt}$$

$$\text{or} \quad \frac{d\phi}{dt} = 2 \frac{d}{dt} (\pi r^2 B) \quad (3.81)$$

This is known as the *betatron condition*, which implies that *in any time interval the linked magnetic flux must change at a rate twice than that would occur if the magnetic field were uniform and equal to the field at the orbit*. This condition makes it necessary to have a central iron core with high flux density inside the orbit. This is obtained by specially contoured pole piece faces where the flux density at the centre of the orbit is greater than it is at the circumference of the orbit. The iron core is laminated, as in a transformer as the induced potential is determined by the rate of change of flux. The betatron condition holds for both non-relativistic and relativistic energies.

The action of a betatron depends on the principle the same as that of a transformer, in which an alternating current applied to the primary induces a proportional and similar current in the secondary coil. The primary current produces an oscillating magnetic field which, in turn, induces an oscillating potential in the secondary coil. The betatron is a transformer, in which a cloud of electrons, located inside an annular, doughnut-shaped vacuum chamber, takes the place of the secondary winding. The chamber is placed within the poles of an electromagnet energized by an alternating pulsed current, and the magnet produces a strong magnetic field in the central space or the doughnut or the hole (Fig. 3.22). The electrons move in a circular orbit of constant radius within the vacuum chamber and they gain energy by electromagnetic induction due to the changing magnetic flux linked with the orbit.

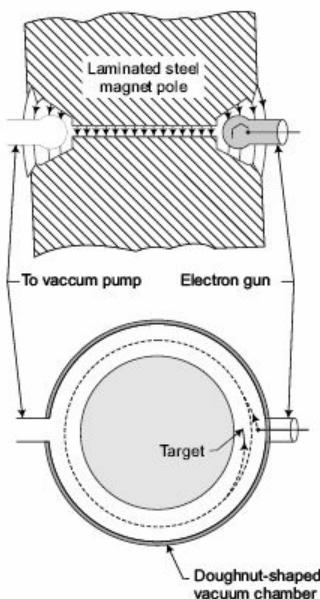


Fig. 3.22 Accelerating chamber in a betatron.

Electrons are injected into the chamber from an electron gun in which thermally generated electrons from a hot cathode are accelerated and focussed by a potential of few thousand volts. The electrons are injected at an instant when the magnetic field is just rising from its zero value in the first quarter cycle, as shown in Fig. 3.23. The increasing magnetic field induces a potential within the doughnut which increases the energy of the electrons. Since the electrons are in a magnetic field, the electrons move in a circular path, and the increasing value of the magnetic field keeps them moving in a stable orbit. When the magnetic field would pass over its maximum value starting to decrease, the direction of the induced magnetic field would reverse and consequently the electrons would slow down. This has to be avoided and therefore the electrons should be removed from their orbit when the value of magnetic field is at peak. For this, a pulse of current is sent through an auxiliary coil which changes the magnetic field in such a way that the electrons are pulled out of their stable orbit and strike a target for producing X-rays or other purposes.

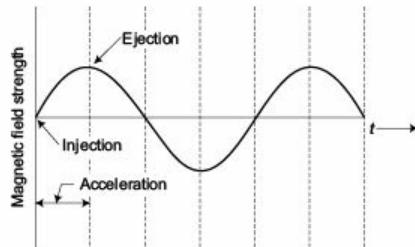


Fig. 3.23 The acceleration and ejection in a betatron.

Electrons with energies of more than 300 MeV have been obtained using betatron.

3.23 NUCLEAR RADIATION DETECTORS

The detection of particles given out from radioactive nuclei or accelerators is based on the interaction of these radiations with matter. In the process of detecting and measuring nuclear particles, there is a sharp distinction between charged and neutral particles. When a charged particle passes through matter (solid, liquid or gas), it can ionize or excite the atoms. This ionization or excitation is detected and forms the basis of the working of the detectors of the charged particles. Neutral particles, such as photons and neutrons, cannot be detected easily. Most of the detectors of the neutral particles work on the principle of the neutral particles producing charged particles (e.g., photon producing photoelectron in the photoelectric effect) and then detecting the charged particle. Different types of detectors have been developed and we discuss below some of them.

3.23.1 Ionization Chamber

An ionization chamber is a detector the working of which is based on the ionization produced by the charged particles. Fig. 3.24 shows a simple ionization chamber which contains two plates, a positively charged plate anode and a negatively charged plate cathode. When the particle or radiation passes through chamber, the gas in the chamber gets ionized and the electrons and the positive ions are produced. The electrons are attracted towards the anode and the positively charged ions towards

the cathode. For each electron reaching the anode, one negative charge flows through the circuit and neutralizes one positive charge at the cathode. In the process a current is produced in the circuit, which can be calibrated to give an idea of the number of particles entering the chamber.

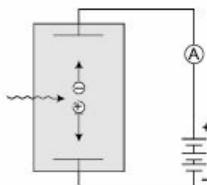


Fig. 3.24 Schematic diagram of an ionization chamber.

The ionization process in the chamber can be multiplied if a high potential difference is applied between the cathode and anode. If the high potential difference is applied, the electrons and the positive ions rush towards the electrodes causing further (secondary) ionization. Figure 3.25 shows the variation of total number of ion collection as a function of applied voltage. If there is no applied voltage across the electrodes, the ions will recombine and almost no current will be produced. As the voltage is increased, for low applied voltage, there is competition between the loss of ion pairs by recombination and the removal of ions by the collection on the electrodes. As the voltage is increased (say to V_1) recombination becomes negligible and the number of ions reaching the electrodes increases. Between the region of applied voltages V_1 and V_2 , in which the number of ion pairs collected is independent of the applied voltage and the curve is horizontal is called the *ionization chamber region*. In this region, as seen in the figure, the number of ions collected n stays constant at 10.

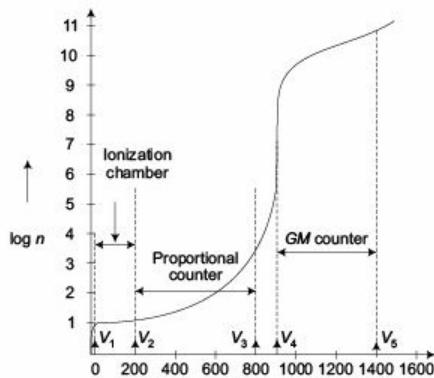


Fig. 3.25 Log of total ion collection ($\log n$) versus applied voltage with region of operation of ionization, proportional and GM counters.

When the voltage is increased above V_2 , n increases above 10 due to a phenomenon called *gas multiplication* or *gas amplification*. At such applied voltages, electrons released in the primary ionization acquire enough energy to produce additional ionization when they collide with gas molecules and n increases almost exponentially with V . Each initial electron produces a small 'avalanche' of electrons with most of the secondary electron ejected close to the electrode. For the some range of voltage, up to V_3 , each electron acts independently and gives its own avalanche, not being affected by the presence of the other electrons. Between V_2 and V_3 , the number of ion pairs collected is proportional to the initial ionization and this is the region of *proportional counter operation*.

Above the voltage V_3 , the gas multiplication effect continues to increase very rapidly and as more electrons produce avalanches. These avalanches interact with one another. The positive ion space charge on one avalanche inhibits the development of the next avalanche. The region between V_3 and V_4 is the region of *limited proportionality*. Above V_4 , the charge collected becomes independent of the ionization initiating it. The gas multiplication increases the total number of ions to a value that is limited by the characteristics of the chamber and the external circuit. The region above V_4 is the region of *Geiger-Muller counter operation*. It ends at the voltage V_5 , where the discharge tends to propagate itself indefinitely. The region above V_5 is the region of *continuous discharge*.

3.24 GEIGER-MULLER (GM) COUNTER

This counter is operated in the applied voltage region (V_4 to V_5) such that the discharge spread throughout the entire length of the counter, resulting in a pulse size independent of the initial ionization. It is especially useful for counting of lightly ionizing particles such as β -particles or γ -rays, although it can also be used for α -particles. It is a portable detector of the nuclear radiations.

A schematic diagram of a GM counter is shown in Fig. 3.26. It consists of a thin (tungsten) wire mounted along the axis of a metallic cylinder. The central wire is positively charged and acts as anode while the metallic cylinder is negatively charged. A gas at a pressure of 2 to 10 cm Hg is filled in the cylinder. The gas is usually a mixture of 90% Argon and 10% ethyl alcohol. A potential difference of 800 to 1000 V is applied to the electrodes. There is a large electric field near the central wire. The whole arrangement is enclosed in glass tube (not shown in the figure).

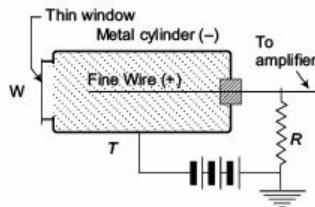


Fig. 3.26 Schematic diagram of a GM counter.

The particles to be counted enter the chamber through a thin foil window W. When a particle enters the chamber, it ionizes an atom. The positive charged ions moves towards the negative electrode. The ejected electrons are drawn towards the positively charged wire, where they are accelerated rapidly producing more ionization. In this process upto 10^9 to 10^{10} ion pairs (or whatever

the counter's maximum capabilities) are produced causing a voltage pulse (\approx of magnitude 10 microvolts). An electronic pulse amplifier amplifies these small voltage pulses in the range of 5 to 50 volts. The amplified voltage pulse is then counted in some way so that the counting rate is measured, using a calibrated scaler.

A typical curve of count rate as a function of tube voltage is shown in Fig. 3.27. When the applied voltage is low, the initial ion pair does not produce a full avalanche although some gas amplification may take place. The size of the current pulse is then proportional to the number of ions initially formed. Consequently, a detector of finite sensitivity will record the large pulses but miss the small ones, resulting in a low count rate. Counters operating in this range can be used to differentiate between different types of particles. With higher tube voltages, the count rate increases and then becomes constant for a range of potential difference. This range is called *plateau* (or *Geiger*) region. For higher applied voltage, one initial particle may give rise to more than one count due to spurious output pulses. This region is of no practical use and can damage the tube.

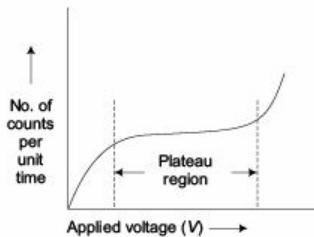


Fig. 3.27 Variation of count rate with potential difference in GM counter.

The counting efficiency of the GM counter is high and all nuclear radiations α -, β - and γ -particles can be detected. It is portable and often used to monitor radiation level around nuclear power plants. It can be used to monitor cosmic rays intensities.

3.25 SOLID STATE DETECTORS

Although it is difficult to replace the gas in an ionization chamber by any solid, but by placing certain semiconductor between the two collecting plates, one can make a detector, known as *solid-state detector*. When energetic particles fall on the semiconductor, they generate pulse with which incident particles can be detected. It is possible to reduce the size of the chamber with the solid state detector.

Such counters have heavily reverse-biased $p-n$ junction diode. When a radiation enters the depletion region, it creates electron-hole pairs. The electron moves toward the n -layer (positive of the battery) and the holes towards the p -type layer (negative of the battery). The electrons so collected form an electronic pulse which is proportional to the energy of the incident radiation. These detectors are smaller in size, require less power and have faster response time.

3.26 SCINTILLATION DETECTORS

When charged particles pass through the matter, they not only ionize atoms but also excite the atoms to higher energy levels. These excited atoms then emit light as they fall back to the ground

state. This emitted light is then exploited in the scintillation detector. Light from a scintillation detector is made to fall on a photoelectric cell which converts the light into an electric pulse, which is then amplified by a photomultiplier.

Scintillation detectors use NaI crystal and certain plastics that are transparent to the light which they produce. This enables the use of a block of material that is thick enough to stop the particles and hence to measure their energy. Some detectors use a liquid scintillator, which allows a radioactive sample to be placed right inside the scintillator, to improve the chance of detecting weak or low-energy signals.

3.27 CLOUD CHAMBER

Photographic film was the most elementary detector, which could produce a visible track of a nuclear particle, even though the particle itself was totally invisible. The cloud chamber detector, invented by C.T.R. Wilson a Scottish physicist in 1910, was an improvement upon the detection by photographic film. In this detector, there is a chamber containing air that is supersaturated with water vapour. When a charged particle passes through a chamber, water droplets condense around the ions produced by the particle. These droplets form a visible trail which can be photographed. If the particle collide with a nucleus in the chamber, any charged particle produced in the nuclear reaction, also produce visible trails. Thus, many details of the nuclear reaction could be revealed using the cloud chamber.

3.28 BUBBLE CHAMBER

It is a variant of cloud chamber and contains superheated liquid hydrogen. The ions formed by any passing charged particle trigger boiling in the nearby liquid, producing tiny bubbles producing a visible record of the particle's path. One of the advantages of bubble chambers is that the density of nuclei in the liquid of the bubble chamber is nearly 1000 times that in the gas of a cloud chamber. Thus, the probability that a passing particle will produce a nuclear reaction in a bubble chamber is much higher than in a cloud chamber.

If a magnetic field is applied to a cloud or bubble chamber, the path of charged particles get curved. It is possible to get an idea of the particle's momentum by measuring the radius of curvature.

Short Answer Type Questions

1. Discuss the basic features of proton-neutron theory.
2. Define the terms giving suitable examples: isotope, isobar, isotope and isomer.
3. How the radius of a nucleus is related to its mass number? Hence, show that the density of nucleons inside the nucleus is the same for all nuclei.
4. What do you mean by 'nuclear spin'? Explain.
5. What is nuclear magneton? Calculate its value.
6. Give three important features of the nuclear force.
7. Define 'nuclear binding energy' and obtain an expression for it.
8. On what factors the stability of the nuclei depend? Explain.
9. Define: decay constant, half-life and mean life for radioactive nuclei
10. Define 'decay rate' and give its unit.

11. Differentiate between natural and artificial radioactivity giving suitable examples.
12. What is 'carbon dating'? Explain.
13. What do you mean by α -decay? Explain.
14. What is positron emission? Give an example.
15. What do you mean by 'electron capture'? Give an example.
16. What is gamma decay? What are isomeric transitions?
17. What is a 'nuclear reaction'? How it differs from a chemical reaction?
18. Define nuclear cross-section and give its unit.
19. Explain why a nucleus can be considered to be analogous to a liquid drop.
20. What are 'nuclear magic numbers'? Why are they called so? What their existence indicate about the nuclear structure?
21. What is 'nuclear fission'? Give an example and explain its significance.
22. Define 'neutron reproduction factor' and give its values for controlled and uncontrolled chain reactions.
23. What are thermal neutrons? What is their significance?
24. What is uranium enrichment? Explain.
25. What is fission decay series? Explain giving an example.
26. Differentiate between prompt and delayed neutrons in the fission of ^{235}U nuclei.
27. For what purposes nuclear reactors are used? Explain.
28. What are moderators? Name any two materials used as moderator.
29. What is 'nuclear fusion'? Why is it called thermonuclear reaction?
30. What are accelerators? Why are they needed?
31. Differentiate between dc and ac accelerators.
32. What do you mean by cyclotron frequency? Explain.
33. What is a 'synchrocyclotron'?
34. What is a betatron? Mention the two differences w.r.t. which it differs from cyclotron.
35. On what principle nuclear radiation detectors work? Explain.
36. What is an ionization chamber? Giving a schematic diagram explain its working.
37. What do you mean by gas multiplication? What is the value of gas multiplication factor for the working region of GM counter?
38. What is working principle of a scintillation detector?
39. Why is a cloud chamber so called?
40. In what respect a bubble chamber differ from a cloud chamber? Explain.

Long Answer Type Questions

1. What is 'nucleus'? Discuss its important features.
2. Discuss the 'magnetic moment' of a nucleus. Hence, define the term 'nuclear magneton'.
3. Describe the important features of nuclear force.
4. Discuss meson theory of nuclear force giving suitable diagram.

5. With the help of a diagram show the variation of binding energy per nucleon with mass number. Discuss its important features.
6. Draw the 'line of stability' and discuss the stability with the variation of atomic number.
7. What is radioactivity? Obtain radioactive decay law and hence define decay constant, half-life and mean life for radioactive nuclei.
8. What is artificial radioactivity? Explain giving suitable example.
9. Discuss various applications of artificial radioactivity.
10. Explain the theory of α -decay. Why half-lives of the alpha particle decaying nuclei vary enormously?
11. What is β -decay? Explain how β -decay could be explained with the help of neutrino.
12. Obtain the expression for Weizsäcker's semiempirical binding energy formula explaining the corrections introduced in the original formula.
13. Explain the process of chain reaction in 'nuclear fission'.
14. Draw fission yield curve for the fission of ^{235}U and discuss its important features.
15. Giving a schematic diagram, explain the working of a nuclear reactor.
16. What is 'nuclear fusion'? Give the condition when it process through $p-p$ cycle and when through $C-N$ cycle. Explain the two cycles.
17. Describe the construction and working of a linear accelerator.
18. Giving a schematic diagram explain the working of a cyclotron.
19. Obtain the betatron condition and explain the working of a betatron.
20. Giving a schematic diagram explain the working of a GM counter.

NUMERICAL EXERCISES

1. Obtain the approximate ratio of the nuclear radii of gold isotope $^{197}_{79}\text{Au}$ and silver isotope $^{107}_{47}\text{Ag}$. What is the approximate ratio of their nuclear densities? (Ans. 1.23, 1)
2. Calculate the binding energy per nucleon in the nuclei of ^{31}P . Given: $m(^{31}\text{P}) = 30.97376u$, $m(^1_0\text{n}) = 1.00865u$, $m(^1_1\text{H}) = 1.00782u$. (Ans. 8.47 MeV)
3. Express 16 mg mass into equivalent energy. (Ans. 9.10^{30} ev)
4. Express one eV in kilowatt-hour. (Ans. 4.4×10^{-26} k Wh)
5. Calculate the minimum energy of γ -rays necessary to disintegrate a deuteron into a proton and a neutron. Given: $m(^2_1\text{H}) = 2.0147034u$, $m(^1_1\text{H}) = 1.007825u$, $m(^1_0\text{n}) = 1.008665u$. (Ans. 1.747 Mev)
6. Calculate the binding energy of a deuteron, from the data given above. (Ans. 2.2 MeV)
7. Calculate the binding energy of an α -particle. Given: $m(^4_2\text{He}) = 4.002602u$.
8. What are the approximate radii of ^{12}C and ^{14}C ? Calculate by what per cent the radii of ^{14}C is higher than that of ^{12}C . (Ans. 2.75 fm, 2.89 fm, 5.27%)
9. What is the binding energy per nucleon for ^{238}U if $m(^{238}\text{U}) = 238.05078u$. (Ans. 7.57 MeV)
10. Show that the radius of $^{238}_{92}\text{U}$ is only 6 times that of ^1_1H .
11. How much energy is required to remove (a) a proton and (b) a neutron from $^{16}_8\text{O}$? Explain the difference between the two energies. (Ans. (a) 12.13 MeV, (b) 15.66 MeV)
12. Find the energies needed to remove a neutron from ^4_2He , then a proton and then finally to separate the remaining neutron and proton. (Ans. 20.6 MeV, 5.5 MeV, 2.2 MeV)

13. The half-life of $^{238}_{92}\text{U}$ against α -decay is 4.5×10^9 years. How many α -decays per second occur in 1 gm of $^{238}_{92}\text{U}$. (Ans. 1.23×10^4 per second)
14. There is a stream of neutrons with a kinetic energy of 0.037 eV. Assuming the half-life neutron to be 700 seconds, what fractions of neutrons will decay before they travel a distance of 10 km. ($m_n = 1.675 \times 10^{-27}$ kg) (Ans. 0.004)
15. The radioactive decay constants of uranium and radon are $4.9 \times 10^{-8} \text{ s}^{-1}$ and $2.1 \times 10^{-6} \text{ s}^{-1}$ respectively. Calculate the mean lives and half-lives of these nuclei. (Ans. 4.5×10^9 yr, 6.5×10^9 yr, 3.8 days, 5.52 days)
16. The half-life of ^{198}Au is 2.7 days. (a) Calculate its decay constant. (b) If at some time, a sample contains one μg of ^{198}Au , what is its activity? (c) How many decays occur after 8 days in this sample? (Ans. (a) $2.97 \times 10^{-6} \text{ s}^{-1}$, (b) 0.244 Ci, (c) 1.16×10^9 decays per second).
17. The half-life of ^{14}C is 5570 years. Find the mass of 2m Ci of radioactive ^{14}C . (Ans. 4.36×10^{-4} g)
18. $^{11}_6\text{C}$ decays to $^{11}_5\text{B}$ by β^+ -emission. What is the maximum energy the neutrino can have? What is its minimum energy? $m(^{11}\text{C}) = 11.0114334 \text{ u}$, $m(^{11}\text{B}) = 11.009305 \text{ u}$. (Ans. 0.96 MeV, zero)
19. The isotope ^{12}N undergo positron emission to ^{12}C which decays to the ground state with the emission of 94.43 MeV gamma ray. Calculate the maximum kinetic energy of the beta particle. ($m(^{12}\text{N}) = 12.018613 \text{ u}$) (Ans. 11.89 MeV)
20. ^{40}K decays to ^{40}Ar by electron capture. Calculate the kinetic energy of the neutrino assuming that the initial kinetic energy of the electron and the recoil energy of the nucleons are zero. (Ans. 1.504 MeV)

Unit IV

4

Electromagnetic Theory

4.1 INTRODUCTION

Visible light, radio and TV signals, signal of mobile phones, X-rays, etc. are all electromagnetic waves. Our daily life is heavily dependent on electromagnetic waves, as from microwave oven at our home to our constant companion, the mobile phones, all use electromagnetic waves for their operation.

The nature of electromagnetic waves could be understood, when the Maxwell's equations combined the electricity and magnetism into one discipline of *electromagnetism*. The Maxwell's equations show that a time-varying magnetic field acts as a source of electric field and a time-varying electric field acts as a source of magnetic field. These **E** and **B** fields can sustain each other forming an electromagnetic waves that propagates through space, without the requirement of any material medium.

In this chapter, we shall use Maxwell's equations as the theoretical basis for understanding electromagnetic waves. We shall see that these waves carry both energy and momentum.

4.2 MOTION OF CHARGED PARTICLES IN ELECTRIC AND MAGNETIC FORCE

Force on a charged particle in an Electric field

When a charged particle having charge q is placed in an electric field **E**, it experiences a force **F**, given by

$$\mathbf{F} = q \mathbf{E} \quad (4.1)$$

The SI unit of force is Newton (N). Thus, if a charged particle of one coulomb is placed in an electric field of 1 N/C, it experience a force of 1 N. If q is positive, the force experienced by the charge is in the same direction as **E** and if q is negative, the force **F** and the field **E** are in opposite directions (Fig. 4.1).

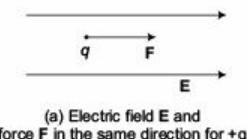
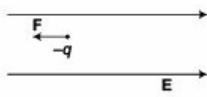
(a) Electric field E and force F in the same direction for $+q$ (b) Electric field E and force F in opposite direction for $-q$

Fig. 4.1

Force on a Charged Particle in a Magnetic Field

When a charged particle having charge q moving with a velocity v enters a region with magnetic field B , it experiences a force F given by

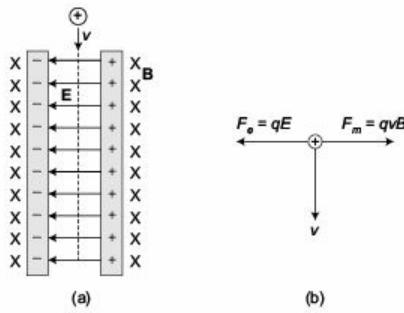
$$F = q v B \sin \theta \quad (4.2)$$

where θ is the angle between the directions of v and B . In vector notation, Eq. (4.2) can be expressed as

$$\mathbf{F} = q (\mathbf{v} \times \mathbf{B}) \quad (4.3)$$

Equation (4.2) gives the magnitude of the force. If $v = 0$, $F = 0$, i.e., a charged particle at rest experiences no magnetic force. This is quite different from the electric field force, in which case force is the same whether the charge is moving or stationary.

If $\theta = 0^\circ$ or 180° , $F = 0$ and $\theta = 90^\circ$, $F = F_{\max}$ i.e., when the charge is moving in the same or opposite direction as the magnetic field, no force acts on the charge due to magnetic field, while the force is maximum when the direction of the velocity is perpendicular to the direction of the magnetic field. It follows from Eq. (4.3) that *the direction of the force F is always perpendicular to both the field B and the velocity v .* The direction of the force is given by right-hand rule.

Fig. 4.2 (a) Velocity selector: perpendicular E and B field, (b) directions of electric force F_e and magnetic force F_m on a positively charged particle.

Lorentz force When a charged particle moves through a region of space, where both electric and magnetic fields are present, both fields exert forces on the particle. The total force \mathbf{F} , also called Lorentz force, is the vector sum of the electric and magnetic forces:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (4.4)$$

Crossed Electric and Magnetic Fields: Velocity Selector

Suppose a charged particle of mass m , charge q and speed v enters a region of space where the electric and magnetic fields are perpendicular to each other (also called crossed electric and magnetic fields) and to the particle's velocity. As shown in Fig. 4.2. The electric field \mathbf{E} is directed towards left and the magnetic field \mathbf{B} into the plane of the paper. If q is positive, the electric force ($= qE$) is towards left and the magnetic force ($= qvB$) is towards right. For given field magnitudes E and B and for a particular value of v , the electric and magnetic forces will be equal in magnitude. In this situation, the net force on the particle is zero and the particle travels in a straight line with constant velocity. That is,

$$qE = qvB$$

or $v = \frac{E}{B}$ (4.5)

Thus, only the particles with speed $v = \frac{E}{B}$ can pass without being deflected by the fields. This

concept is used to design *velocity selector*, in which the particles of a specific speed can be selected from a beam of particles with the arrangement of crossed electric and magnetic fields. By adjusting the value of E and B , the particles with any speed v can be selected. The Eq. (4.5) is also valid for the negatively charged particles. In this case, the direction of electric force and magnetic force gets reversed.

Example 4.1

In Thomson's *e/m* experiment, the electrons are accelerated at an accelerating potential of 150 V and a deflecting electric field of magnitude 6.0×10^6 N/C. (a) At what fraction of the speed of light the electrons are moving? (b) What magnitude of the magnetic field will be needed for the velocity selector? (c) With this magnetic field, what will happen if the accelerating potential is above 15 V?

Solution. (a) Here, $V = 150$ volts, $\frac{e}{m}$ for electron = 1.76×10^{11} C/kg

We know $\frac{1}{2}mv^2 = eV$

or $v = \sqrt{2 \cdot \frac{e}{m} \cdot V}$

$$\begin{aligned} &= \sqrt{2 \times 1.76 \times 10^{11} \text{ C/kg} \times 150 \text{ V}} \\ &= 7.27 \times 10^6 \text{ m/s} \\ \frac{v}{c} &= \frac{7.27 \times 10^6 \text{ m/s}}{3 \times 10^8 \text{ m/s}} = 0.024 \end{aligned}$$

(b) As

$$v = \frac{E}{B}$$

$$\therefore B = \frac{E}{B} = \frac{6.00 \times 10^6 \text{ N/C}}{7.27 \times 10^6 \text{ m/s}} = 0.83 \text{ T}$$

- (c) Increasing the accelerating potential increases the speed v of the electrons. As a consequence, the force due to the magnetic field ($= e v B$) will increase but the force due to the electric field ($= e E$) will remain unchanged. Therefore, the electron beam will be deflected in the direction of the magnetic force.

4.3 MAGNETIC FOCUSsing

Magnetic field can be used to focus, i.e., to converge electrons. This is used in electron microscopes as a *magnetic lens*.

Suppose an electron or any charged particle moving with velocity v enters a region of uniform magnetic field B making an angle θ with the direction of the field (Fig. 4.3). Resolving the velocity v , represented by OP , into the two components, $v \cos \theta$ is the component in the direction of the field while $v \sin \theta$ is the component perpendicular to the direction of the field. If only the parallel component $v \cos \theta$ is present, the electron will move in the direction of the magnetic field. While in the presence of only the perpendicular component $v \sin \theta$, the electron will move in a circular path, when both the components are present, *the resultant path of the electron is a helix*.

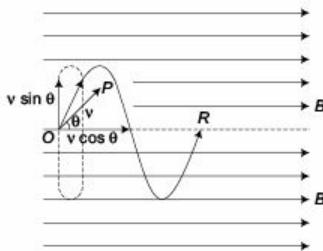


Fig. 4.3 Magnetic focussing

The radius r of the circle due to the component $v \sin \theta$ is given by:

$$\frac{m(v \sin \theta)^2}{r} = Be v \sin \theta \quad (4.6)$$

where, m is the mass of the electron and e the charge on it.

$$\therefore v \sin \theta = \frac{Ber}{m}$$

Time taken by the electron to complete the circle:

$$t = \frac{2\pi r}{v \sin \theta} = \frac{2\pi m}{Ber}$$

$$= \frac{2\pi m}{Be} = \frac{2\pi}{B\left(\frac{e}{m}\right)} \quad (4.7)$$

The pitch of the helix l is the distance travelled by the electron parallel to the magnetic field in time t .

$$\begin{aligned} l &= v \cos\theta \times t \\ &= v \cos\theta \times \frac{2\pi m}{Be} = \frac{2\pi m v \cos\theta}{Be} \end{aligned} \quad (4.8)$$

It is clear from Eq. (4.7) that the time t is independent of angle θ . Therefore, an electron starting from O in any direction will arrive at R at the same time. In the absence of the magnetic field, the electron would have moved along OP . Thus, it follows that if electrons emitted from a source at O and forming a divergent beam, is subjected to a uniform magnetic field, converge at a point R , distant l , equal to the pitch of the helix. The value of l can be altered by changing the value of the applied magnetic field B . The action of the magnetic field is thus similar to that of a converging lens on a divergent beam of light rays. Hence, the magnetic field is referred to as *magnetic lens* of focal length l .

If l is so adjusted that the electrons make two or more complete rotations in the helix, the focal point will occur at $2l, 3l, \dots$ Thus, unlike an optical lens which has a fixed focal length, a magnetic lens has a variable focal length determined by the value of applied magnetic field.

4.4 GAUSS'S LAW

Gauss's law gives a connection between electric flux (Φ_E) of any closed hypothetical surface (called a Gaussian surface) and the net charge enclosed by the surface. It says that *the flux of the electric field through any closed surface equals the total net charge inside the enclosed volume divided by ϵ_0* . Mathematically,

$$\Phi_E = \oint \mathbf{E} \cdot d\mathbf{S} = \frac{q}{\epsilon_0} \quad (4.9)$$

where q is the total charge inside the volume enclosed by the surface. If a surface encloses equal and opposite charge, the flux is zero. Charge outside the closed surface makes no contribution to the value of q , nor does the exact location of the q inside affect this value of charge and hence the flux.

Suppose a single positive point charge q , radiating electric field lines equally in all directions, is located at the centre of an imaginary spherical surface with radius r (Fig. 4.4). The magnitude of the electric field at every point on the surface is given by

$$E = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \quad (4.10)$$

At each point on the surface, E is perpendicular to the surface and its magnitude is the same at every point. The total electric flux Φ_E is given by

$$\Phi_E = E \cdot A = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \cdot 4\pi r^2 = \frac{q}{\epsilon_0} \quad (4.11)$$

This is the Gauss' law. It is clear from Eq. (4.11), *the flux is independent of the radius of the sphere, it depends only on the charge enclosed by the sphere*. If q is negative

$$\Phi_E = -\frac{q}{\epsilon_0} \quad (4.12)$$

Suppose the total charge inside a Gaussian surface S is zero. Obviously, the total flux Φ_E through S is also zero. A line of force entering S makes a negative contribution to flux, while those leaving S make positive contribution to flux. So if the total enclosed charge is zero and S is placed in an electric field there must be same number of lines entering S and leaving S .

Example 4.2

A point charge of 10×10^{-6} Coulomb is at the centre of a cubical Gaussian surface of sides 0.5 m. Calculate flux for the surface.

Solution. We know that for a Gaussian surface

$$\Phi_E = \frac{q}{\epsilon_0}$$

Here,

$$q = 1.0 \times 10^{-6} \text{ Coulomb}$$

Taking,

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ Coulomb}^2/\text{Newton-m}^2$$

$$\begin{aligned}\Phi_E &= \frac{1.0 \times 10^{-6}}{8.85 \times 10^{-12}} \frac{\text{N-m}^2}{\text{C}} \\ &= 1.12 \times 10^5 \frac{\text{N-m}^2}{\text{C}}\end{aligned}$$

Coulomb's law from Gauss's law: Coulomb's law can be deduced from Gauss's law and symmetry considerations. Let us apply Gauss' law to an isolated point charge q (Fig. 4.4). Although Gauss' law holds for any surface it is much easier to extract information for a spherical surface of radius r centred on the charge. The advantage of this surface is that, from symmetry, E must be normal to it and must have the same magnitude for all points on the surface.

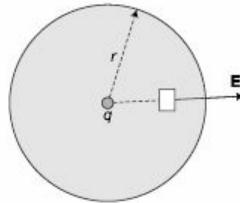


Fig. 4.4 A spherical Gaussian surface of radius r surrounding a point charge q .

In Fig. 4.4, both E and dS at any point on the Gaussian surface are directed radially outward. The angle between them is zero and the quantity $E \cdot dS$ becomes simply EdS . Gauss' law thus becomes

or, $\Phi_E = \oint \mathbf{E} \cdot d\mathbf{S} = \frac{q}{\epsilon_0}$

or, $\epsilon_0 E \oint dS = q \quad (\text{as } E \text{ is constant at all points on the sphere})$

or, $\epsilon_0 E \cdot 4\pi r^2 = q$

or, $E = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \quad (4.13)$

This equation gives the magnitude of the electric field \mathbf{E} at a distance r from q . The direction of \mathbf{E} can be determined from symmetry.

Say, a point charge q_0 is placed at the point at which \mathbf{E} is calculated the magnitude of the force that acts on the test charge is

$$F = E q_0$$

Putting value of E from Eq. (4.13), we get

$$F = \frac{1}{4\pi\epsilon_0} \frac{q q_0}{r^2} \quad (4.14)$$

which is Coulomb's law.

Gauss' law is one of the four fundamental equations of electromagnetic theory.

4.5 APPLICATIONS OF GAUSS'S LAW

Gauss's law can be used to calculate electric field if symmetry exists in the charge distribution so that a Gaussian surface can be visualized.

We have already calculated electric field due to a point charge in the previous section while deriving Coulomb's law from Gauss's law. Now we calculate \mathbf{E} due to some continuous charge distributions.

4.5.1 Field Due to a Straight Line of Charge

We calculate the field due to an infinite line of charge, having a uniform linear charge density λ at all points on the line (Fig. 4.5). Let us find an expression for \mathbf{E} at a distance r from the line.

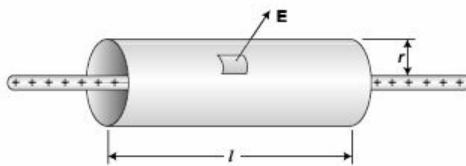


Fig. 4.5 An infinite line of charge, indicating a cylindrical Gaussian surface.

We can easily visualize from the symmetry considerations that the electric field is radially directed and is independent of both positions along the wire and angular position around the wire. Obviously, locus of equal magnitude electric field will be a cylinder co-axial with the linear charge distribution.

That is, we choose a right cylinder of radius r and length l as the closed Gaussian surface.

Since the electric field is parallel to the circular ends, i.e., θ angle between \mathbf{E} and $d\mathbf{S}$, equals 90° , contribution of these ends to the flux integral is zero ($\therefore \mathbf{E} \cdot d\mathbf{S} = 0$). The only contribution to the flux integral comes from the curved surface of the cylinder where $\theta = 0^\circ$. The charge enclosed within Gaussian surface is λl . using Gauss's law.

$$\oint \mathbf{E} \cdot d\mathbf{S} = \frac{q}{\epsilon_0}$$

$$E \cdot 2\pi r l = \frac{\lambda l}{\epsilon_0}$$

$$E(r) = \frac{\lambda}{2\pi\epsilon_0 r}$$

$$\text{In vector form, } \mathbf{E}(r) = \frac{\lambda}{2\pi\epsilon_0 r} \hat{\mathbf{r}} \quad (\hat{\mathbf{r}} - \text{a unit vector in the radial direction}) \quad (4.15)$$

For a line of positive charge, \mathbf{E} is radially outward, while that for negative charge radially inward.

4.5.2 Field Due to an Infinite Sheet of Charge

Let us consider a thin plane infinite sheet of positive charge, having a uniform surface charge density σ (Fig. 4.6). From symmetry considerations it is clear that the electric field must be perpendicular to the plane sheet. Equidistant points on the opposite sides must have \mathbf{E} equal and oppositely directed. In this case a pillbox of cross-sectional area A and length $2r$ is chosen as the Gaussian surface. The two circular end surfaces are the surfaces which will contribute to the flux integral, as $\theta = 0^\circ$ for these surfaces. Applying Gauss's law, we get

$$\oint \mathbf{E} \cdot d\mathbf{S} = \frac{q}{\epsilon_0}$$

Here, $q = \text{Surface charge density} \times \text{Surface area}$

$$= \sigma A$$

$$\therefore 2 E A = \frac{\sigma A}{\epsilon_0}$$

(as there are two surfaces of area A , $\therefore \int dS = 2A$)

$$\mathbf{E} = \frac{\sigma}{2\epsilon_0} \quad (4.16)$$

We observe that E is independent of the distance of the point r or E is the same for all points on both sides of the plane.

If charge distributed is negative, then

$$E = -\frac{\sigma}{2\epsilon_0} \quad (4.17)$$

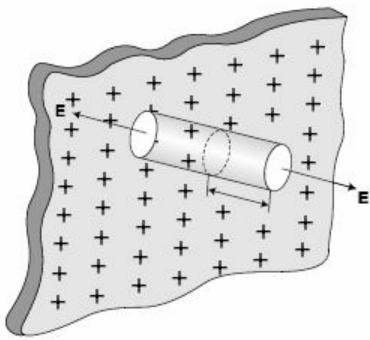


Fig. 4.6 An infinite sheet of positive charge pierced by a cylindrical Gaussian surface.

In vector form, if $\hat{\mathbf{r}}$ is unit vector perpendicular to the sheet, E is expressed as

$$\mathbf{E} = \frac{\sigma}{2\epsilon_0} \hat{\mathbf{r}}, \text{ towards right of the sheet} \quad (4.18)$$

$$\mathbf{E} = -\frac{\sigma}{2\epsilon_0} \hat{\mathbf{r}}, \text{ towards left of the sheet} \quad (4.19)$$

4.5.3 Field Due to Two Infinite Parallel Sheets of Charge

Let us consider two thin infinite parallel plane sheets of charge (Fig. 4.7). Say, σ_1 and σ_2 (Cm^{-2}) be uniform charge densities on the two sheets and $\sigma_1 > \sigma_2 > 0$. There are three regions and in each the field has a constant magnitude and direction. Three regions are indicated in Fig. 4.7. We can find \mathbf{E} in three regions using Eqs. (4.18) and (4.19) and the superposition principle, that is, adding the fields separately produced by each sheet. Therefore, field in each region is

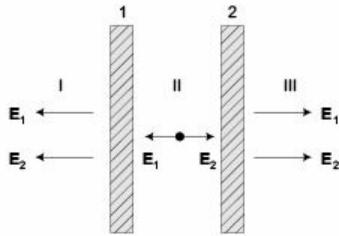


Fig. 4.7 Field due to two infinite plane parallel sheets of charge.

Region I:

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_1 + \mathbf{E}_2 \\ &= -\frac{\sigma_1}{2\epsilon_0} \hat{\mathbf{r}} - \frac{\sigma_2}{2\epsilon_0} \hat{\mathbf{r}} \\ &= -\frac{1}{2\epsilon_0} (\sigma_1 + \sigma_2) \hat{\mathbf{r}} \quad \left(\text{or, } E = -\frac{1}{2\epsilon_0} (\sigma_1 + \sigma_2) \right) \end{aligned} \quad (4.20)$$

Region II:

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_1 + \mathbf{E}_2 \\ &= \frac{\sigma_1}{2\epsilon_0} \hat{\mathbf{r}} - \frac{\sigma_2}{2\epsilon_0} \hat{\mathbf{r}} \\ &= \frac{1}{2\epsilon_0} (\sigma_1 - \sigma_2) \hat{\mathbf{r}} \quad \left(\text{or, } E = \frac{1}{2\epsilon_0} (\sigma_1 - \sigma_2) \right) \end{aligned} \quad (4.21)$$

Region III:

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_1 + \mathbf{E}_2 \\ &= \frac{\sigma_1}{2\epsilon_0} \hat{\mathbf{r}} + \frac{\sigma_2}{2\epsilon_0} \hat{\mathbf{r}} \\ &= \frac{1}{2\epsilon_0} (\sigma_1 + \sigma_2) \hat{\mathbf{r}} \quad \left(\text{or, } E = \frac{1}{2\epsilon_0} (\sigma_1 + \sigma_2) \right) \end{aligned} \quad (4.22)$$

When $\sigma_1 = \sigma$ and $\sigma_2 = -\sigma$, i.e., two infinite plane parallel sheets with equal and opposite uniform charge densities, we see:

from Eq. (4.20), in region I, $\mathbf{E} = 0$

from Eq. (4.21), in region II, $E = \frac{\sigma}{\epsilon_0}$ (a constant)

from Eq. (4.22), in region III, $\mathbf{E} = 0$

So we find that the field is non-zero only between the charged sheets. The field is constant and points from the positive sheet to the negative sheet. In the regions I and III, the field due to two sheets are equal opposite and cancel each other.

This is a practical way of producing a uniform electric field.

4.5.4 Field Due to Uniformly Charged Spherical Shell

Consider a thin spherical shell of radius R with a uniform surface charge density $\sigma \text{ Cm}^{-2}$ (Fig. 4.8). The suitable Gaussian surface is a concentric spherical surface since on such a surface the magnitude of the field is the same at all points and is radially outward directed. In case charge is negative, it will be radially inward.

The magnitude of the field can only depend on distance r of the point P from the centre of the shell, so we write $E(r)$. Considering a point P lying outside the hollow sphere and applying Gauss' law to the Gaussian spherical surface of radius r ($r > R$), we get

$$\int \mathbf{E} \cdot d\mathbf{s} = \frac{q}{\epsilon_0}$$

$$E(r) \cdot 4\pi r^2 = \frac{4\pi R^2 \sigma}{\epsilon_0}$$

$$E(r) = \frac{\sigma R^2}{\epsilon_0 r^2} \quad (4.23)$$

If charge distributed is negative then

$$E(r) = -\frac{\sigma R^2}{\epsilon_0 r^2} \quad (4.24)$$

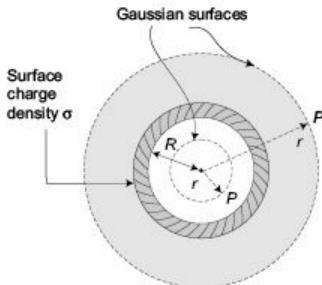


Fig. 4.8 A spherical shell of radius R .

Using $q = 4\pi R^2 \sigma$, Eq. (4.23) can be written as

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

or

$$\mathbf{E}(r) = \frac{q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}} = \frac{q}{4\pi\epsilon_0 r^3} \mathbf{r} \quad (4.25)$$

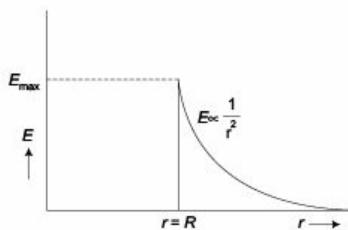
For a point P , inside the shell ($r < R$), charge inside the Gaussian surface is zero and therefore

$$E = 0 \text{ (for } r < R\text{)}$$

From Eqn. (4.25), we can conclude that for points outside a spherically symmetric charged shell, the field is as though the entire charge q is concentrated at the centre of the shell.

The variation of electric field with distance from the centre of a uniformly charged spherical shell is shown in Fig. 4.9.

The fact that the field inside a uniformly charged spherical shell is zero is used for what is called *electrostatic shielding* to protect some very sensitive devices from electric field.

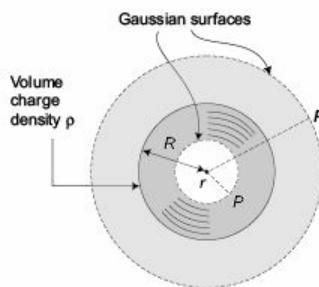
Fig. 4.9. Variation of E with r .

4.5.5 Field Due to a Charged Solid Sphere

Let us consider a uniformly charged solid sphere of radius R (Fig. 4.10). Let q be the total charge uniformly distributed throughout its volume, so that the volume density of charge is given by

$$\rho = \frac{q}{\frac{4}{3} \pi R^3} \text{ C/m}^3$$

only on the radial distance r from the centre. Here also, the suitable Gaussian surface is a concentric spherical surface, say, of radius r . The total flux through spherical Gaussian surface S of radius r is equal to $4\pi r^2 E(r)$. Therefore,

Fig. 4.10 A charged solid sphere of radius R .

$$\oint_S \mathbf{E} \cdot d\mathbf{S} = \frac{q}{\epsilon_0}$$

$$4\pi r^2 E(r) = \frac{q}{\epsilon_0} \Rightarrow E(r) = \frac{q}{4\pi\epsilon_0 r^2} \quad (4.26)$$

Using $q = \rho \cdot \frac{4}{3} \pi R^3$, we get

$$E(r) = \frac{\rho}{3\epsilon_0} \frac{R^3}{r^2} \quad (r \geq R)$$

In the vector form

$$\mathbf{E}(r) = \frac{qR^3}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}$$

or,

$$\begin{aligned} \mathbf{E}(r) &= \frac{\rho}{3\epsilon_0} R^3 \frac{\hat{\mathbf{r}}}{r^3} \\ &= \frac{\rho R^3}{3\epsilon_0 r^2} \hat{\mathbf{r}} \end{aligned} \quad (4.27)$$

If the Gaussian surface S happens to be inside ($r < R$) the solid sphere, the charge enclosed is equal to

$$q' = \frac{4}{3}\pi r^3 \times \rho$$

Therefore,

$$4\pi r^2 E(r) = \frac{\rho \cdot \frac{4}{3}\pi r^3}{\epsilon_0}$$

or,

$$E(r) = \frac{\rho}{3\epsilon_0} r \quad (4.28)$$

or,

$$\mathbf{E}(r) = \frac{\rho}{3\epsilon_0} \mathbf{r} \quad (4.29)$$

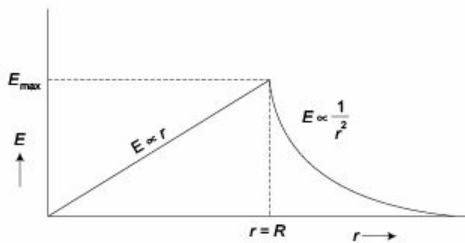
Clearly, $E(r) \propto r$, i.e., electric field increases as r increases.

As at the centre of sphere, $r = 0$, there, $E = 0$.

From Eqn. (4.26), $E(r) \propto \frac{1}{r^2}$ (for $r > R$), the field behaves as if all charges were concentrated at the centre of sphere.

At the surface of sphere $r = R$, the field is maximum, given by

$$\begin{aligned} E_{\max} &= \frac{r}{3\epsilon_0} \frac{R^3}{R^2} \\ &= \frac{\rho}{3\epsilon_0} R \end{aligned} \quad (4.30)$$

Fig. 4.11 Variation of E with r .

These variations of electric field E with distance r from the centre of the sphere are shown in Fig. 4.11.

In all the expressions for the electric field, we will have to put a -ve sign in case the charge distributed happens to be negative.

Example 4.3

Two large metal plates of area 5.0 m^2 face each other. They are 5.0 cm apart and carry equal and opposite charges on their inner surface. If E between the plate is 55 N/C , what is the charges on the plates? Neglect edge effects, that is, assume the plates to be infinite.

Solution. We know that the electric field intensity E between the plates is given by

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

and the charge density

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

∴

$$q = A E \epsilon_0$$

Here,

$$E = 55 \text{ N/C}$$

$$A = 5.0 \text{ m}^2$$

and taking

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ C/N-m}^2, \text{ we get}$$

$$q = 5.0 \times 55 \times 8.85 \times 10^{-12} \text{ Coulomb}$$

$$= 2.447 \times 10^{-9} \text{ Coulomb}$$

Example 4.4

A particle of mass $5 \times 10^{-9} \text{ kg}$ is kept over a large horizontal sheet of charge of density $4.0 \times 10^{-6} \text{ C/m}^2$ (Fig. 4.12). What charge should be given to this particle so that if released, it does not fall? How many electrons are to be removed to give this charge? How much mass is decreased due to the removal of these electrons?

Solution. The electric field in front of the sheet is

$$E = \frac{\sigma}{2 \epsilon_0} = \frac{4.0 \times 10^{-6} \text{ C/m}^2}{2 \times 8.85 \times 10^{-12} \text{ C}^2/\text{N-m}^2}$$

$$= 2.26 \times 10^5 \text{ N/C}$$



Fig. 4.12

If a charge q is given to the particle, the electric force qE acts in the upward direction. It will balance the weight of the particle if

$$\begin{aligned} q \times 2.26 \times 10^5 \text{ N/C} &= 5 \times 10^{-9} \text{ kg} \times 9.8 \text{ m/s}^2 \\ q &= \frac{4.9 \times 10^{-8}}{2.26 \times 10^5} \text{ C} \\ &= 2.21 \times 10^{-13} \text{ C} \end{aligned}$$

The charge on one electron is $1.6 \times 10^{-19} \text{ C}$. The number of electrons to be removed

$$= \frac{2.21 \times 10^{-13} \text{ C}}{1.6 \times 10^{-19} \text{ C}} = 1.4 \times 10^6$$

Mass decreased due to the removal of these electrons

$$\begin{aligned} &= 1.4 \times 10^{-6} \times 9.1 \times 10^{-31} \text{ kg} \\ &= 1.3 \times 10^{-24} \text{ kg.} \end{aligned}$$

Example 4.5

A charge of $4 \times 10^{-8} \text{ C}$ is distributed uniformly on the surface of a sphere of radius 1 cm. It is covered by a concentric, hollow conducting sphere of radius 5 cm.

- (i) Find the electric field at a point 2 cm away from the centre.
- (ii) A charge of $6 \times 10^{-8} \text{ C}$ is placed on the hollow sphere. Find the surface charge density on the outer surface of the hollow sphere.

Solution. (i) By symmetry, the electric field at every point of the Gaussian spherical surface drawn through a point P (at a distance of 2 cm from the centre of the charged sphere) is equal and directed radially.

∴ Applying Gauss's theorem over the Gaussian surface, we have

$$\begin{aligned} \oint \mathbf{E} \cdot d\mathbf{S} &= \frac{q}{\epsilon_0} \\ \Rightarrow \oint E dS &= \frac{q}{\epsilon_0} \\ \Rightarrow E \oint dS &= \frac{q}{\epsilon_0} \\ \Rightarrow E \times 4 \pi r^2 &= \frac{q}{\epsilon_0} \end{aligned}$$

where r = radius of Gaussian surface = 2 cm.

$$\begin{aligned} q &= \text{charge on the sphere} \\ &= 4 \times 10^{-8} \text{ C} \end{aligned}$$

$$E = \text{electric field at any point of Gaussian surface}$$

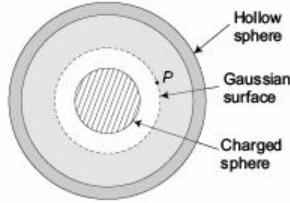


Fig. 4.13

$$\Rightarrow E = \frac{1}{4\pi\epsilon_0} \cdot \frac{q}{r^2}$$

$$= \left(9 \times 10^9 \frac{\text{Nm}^2}{\text{C}^2}\right) \left(\frac{4 \times 10^{-8} \text{C}}{4 \times 10^{-4} \text{m}^2}\right)$$

$$= 9 \times 10^5 \text{ N/C}$$

- (ii) If we imaging a Gaussian surface through the material, of the hollow sphere and apply the Gauss's theorem, then

$$\oint \mathbf{E} \cdot d\mathbf{S} = \frac{q}{\epsilon_0}$$

$$= 0,$$

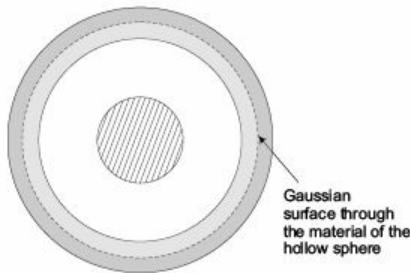


Fig. 4.14

as electric field inside a conductor is zero.

\therefore Net charge enclosed by the Gaussian surface is zero. Thus, the inner surface of the hollow sphere must have a charge of $-4 \times 10^{-8} \text{ C}$. Thus, the charge on the outer surface must be $+4 \times 10^{-8} \text{ C}$.

But the additional charge given on the hollow sphere is $+6 \times 10^{-8} \text{ C}$, the net charge on the outer surface will appear as $+10 \times 10^{-8} \text{ C}$.

\therefore Surface charge density on the outer surface of hollow sphere

$$= \frac{+10 \times 10^{-8} \text{ C}}{4\pi(5 \times 10^{-2} \text{ m})^2}$$

$$= \frac{1}{2\pi} \times 10^{-4} \text{ C/m}^2$$

Example 4.6

The electric field in a region is radially outward with magnitude $E = A x$. Calculate the charge contained in a sphere of radius 20 cm from the origin. Take $A = 100 \text{ N/C-m}$

Solution. Since, the electric fields are radially outward, they are normal to the spherical Gaussian surface with the centre at the origin.

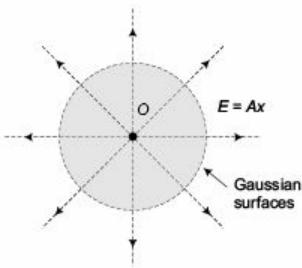


Fig. 4.15

Applying Gauss's theorem

$$\begin{aligned}
 \oint \mathbf{E} \cdot d\mathbf{S} &= \frac{q}{\epsilon_0} \\
 \Rightarrow E \oint dS &= \frac{q}{\epsilon_0} \\
 \Rightarrow E \times 4\pi r^2 &= \frac{q}{\epsilon_0} \\
 \Rightarrow q &= E \times 4\pi\epsilon_0 \times r^2 \\
 &= A r \times 4\pi\epsilon_0 \times r^2 \\
 &= A r^3 \times 4\pi\epsilon_0 \\
 &= \left(100 \frac{\text{N}}{\text{C}\cdot\text{m}}\right) (0.2 \text{ m})^3 \times \left(\frac{1}{9 \times 10^9 \text{ Nm}^2/\text{C}^2}\right) \\
 &= 8.89 \times 10^{-11} \text{ C} \\
 &= \text{charge enclosed by the surface.}
 \end{aligned}$$

Example 4.7

Two parallel metal plates, each with an area A , carry charges Q_1 and Q_2 . The distance between the plates is much less than their linear dimensions. Determine the intensity of the electric field at points A , B and C .

Solution. Electric field due to a large sheet is

$$E = \frac{\sigma}{2\epsilon_0}$$

Let E_A , E_B , E_C be the electric fields at A , B and C respectively. Let us represent the electric field due to plates having charges Q_1 and Q_2 by E_1 and E_2 respectively.

If the direction of the fields be considered negative towards left and positive towards right, then

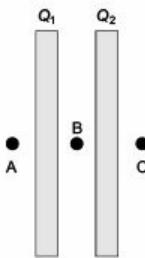


Fig. 4.16

$$\begin{aligned}\mathbf{E}_A &= \mathbf{E}_1 - \mathbf{E}_2 \\ &= -\left(\frac{\sigma_1}{2\epsilon_0} + \frac{\sigma_2}{2\epsilon_0}\right) \\ &= -\left(\frac{Q_1}{2A\epsilon_0} + \frac{Q_2}{2A\epsilon_0}\right) \\ &= -\frac{1}{2A\epsilon_0}(Q_1 + Q_2)\end{aligned}$$

$$\begin{aligned}\mathbf{E}_B &= \mathbf{E}_1 - \mathbf{E}_2 \\ &= \frac{1}{2A\epsilon_0}(Q_1 - Q_2)\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}_C &= \mathbf{E}_1 + \mathbf{E}_2 \\ &= \frac{1}{2A\epsilon_0}(Q_1 + Q_2).\end{aligned}$$

Example 4.8

In Example 4.7, find the charge developed on the four faces of the two plates.

Solution. Considering a Gaussian surface as shown in the diagram, the electric flux through the surfaces lying inside the plates will be zero as there is no electric field inside the material of the plates.

Also, the electric flux through the remaining portion of the surface is zero as the electric field is parallel to them.

Thus, the net flux over the Gaussian surface is zero.

Thus, the inner faces of the plates must carry equal and opposite charge to make the net charge enclosed by the Gaussian surface to be zero.

Let, the equal and opposite charges be $+q$ and $-q$ at the innerfaces of the plates as shown in Fig. 4.17. Thus, the charges on the outer faces would be $Q_1 - q$ and $Q_2 + q$.

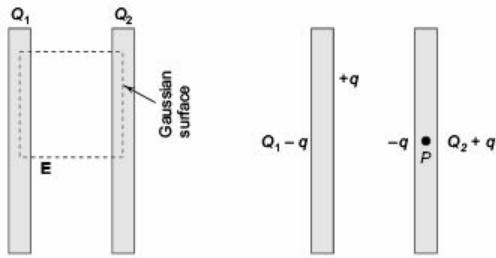


Fig. 4.17.

Now, let us calculate the electric field at P inside a plate.

$$\text{Field at } P \text{ due to } Q_1 - q = \frac{Q_1 - q}{2A\epsilon_0} \text{ (towards right)}$$

$$\text{Field at } P \text{ due to } +q = \frac{q}{2A\epsilon_0} \text{ (towards right)}$$

$$\text{Field at } P \text{ due to } -q = \frac{-q}{2A\epsilon_0} \text{ (towards left)}$$

$$\text{Field at } P \text{ due to } Q_2 + q = \frac{Q_2 + q}{2A\epsilon_0} \text{ (towards left)}$$

and the net field P should be zero.

$$\therefore \frac{Q_1 - q}{2A\epsilon_0} + \frac{q}{2A\epsilon_0} - \frac{-q}{2A\epsilon_0} - \frac{Q_2 + q}{2A\epsilon_0} = 0.$$

$$\Rightarrow Q_1 - q - Q_2 - q = 0$$

$$\Rightarrow q = \frac{Q_1 - Q_2}{2}$$

$$\therefore \text{Charge on left face of left plate} = Q_1 - q$$

$$= \frac{Q_1 + Q_2}{2}$$

$$\text{Charge on right face of left plate} = q$$

$$= \frac{Q_1 - Q_2}{2}$$

$$\text{Charge on left face of right plate} = Q_1 - q$$

$$= \frac{Q_2 - Q_1}{2}$$

and charge on right face of right plate

$$\begin{aligned} &= Q_2 + q \\ &= \frac{Q_1 + Q_2}{2} \end{aligned}$$

4.6 MAXWELL'S DISPLACEMENT CURRENT

Inconsistency of Ampere's law: Ampere's circuital law, relating magnetic field B and electric current I is given by

$$\oint_c \mathbf{B} \cdot d\mathbf{l} = \mu_0 I \quad (4.31)$$

Here, I is the conduction current passing through the surface bounded by closed path C . This law reveals that an electric current produces magnetic field and gives a method to calculate the magnetic field.

Maxwell in 1864, showed that Ampere's law in the form expressed by Eq. (4.19) is logically inconsistent. To understand the inconsistency, let us consider the following situation.

Consider a parallel plate capacitor (Fig. 4.18) connected across battery. Let I be the current due to charging of capacitor. Ampere's law is applied to two loops C_1 and C_2 . As shown in the figure, C_1 encircles the wire connected to the plate P of the capacitor and C_2 encircles the region between the two plates of the capacitor.

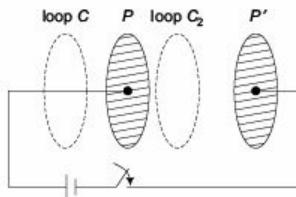


Fig. 4.18

Clearly, a current I is flowing across C_1 while there is no current through C_2 , since there is no wire between the plates through which the conduction current can flow. Here Ampere's law for two paths C_1 and C_2 gives

$$\oint_{C_1} \mathbf{B} \cdot d\mathbf{l} = \mu_0 I \quad (4.32)$$

$$\oint_{C_2} \mathbf{B} \cdot d\mathbf{l} = 0 \quad (4.33)$$

The Eqs. (4.32) and (4.33) should continue to hold even if C_1 and C_2 are infinitesimally close to the plate P of the capacitor. However, if C_1 and C_2 are infinitesimally close to each other, then it is expected that

$$\oint_{C_1} \mathbf{B} \cdot d\mathbf{l} = \oint_{C_2} \mathbf{B} \cdot d\mathbf{l}$$

But this is clearly contradictory to Eqs. (4.32) and (4.33). Because, this contradiction arise due to use of Ampere's circuital law, Maxwell felt that something is missing in this law, that is Ampere's circuital law is inconsistent.

Displacement Current. To remove the contradiction, Maxwell modified Ampere's circuital law. The modified Ampere's circuital law is given by

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \left(I + \epsilon_0 \frac{d\Phi_E}{dt} \right) \quad (4.34)$$

where Φ_E is the electric flux.

Also

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 (I + I_D) \quad (4.35)$$

where

$$I_D = \epsilon_0 \frac{d\Phi_E}{dt} \quad (4.36)$$

The conduction current is denoted by I while I_D is the displacement current and exists as long as electric field (or flux) changes with time, i.e., $\frac{d\Phi_E}{dt} \neq 0$. Thus, the displacement current exists in the region where electric flux is changing with time. The modified law is often called Ampere-Maxwell law.

The modification (or generalization) of Ampere's law implies that the source of magnetic field is not only the conduction current I due to flowing charges but also the changing electric field. In other words, the total current is the sum of conduction current and displacement current. In case of a charging capacitor, both conduction and displacement currents are present. In certain region of space there may be no conduction current but only displacement current. In such a region, a magnetic field exists, although there is no (conduction) current source nearby. This can be easily seen by observing a magnetic field in between the plates of a capacitor.

Example 4.9

Show that when a parallel-plate capacitor is being charged the displacement current across the area in the region between the plates and parallel to \mathbf{E} (Fig. 4.18) is equal to conduction current in the conduction wires.

Solution. We know the displacement current is

$$I_D = \epsilon_0 \frac{d\Phi_E}{dt}$$

The electric field between the plates, if A is its area of cross-section and Q the accumulated charge on each plate, is given by

$$E = \frac{Q}{\epsilon_0 A}$$

$$\therefore \text{Electric flux through area } A, \quad \Phi_E = E \times A = \frac{Q}{\epsilon_0 A} \times A = \frac{Q}{\epsilon_0}$$

$$\therefore I_D = \epsilon_0 \frac{d}{dt} \left(\frac{Q}{\epsilon_0} \right) = \frac{dQ}{dt} = I$$

where $\frac{dQ}{dt}$ is the rate of flow of charge through the connecting wire which is equal to the conduction current.

4.7 CONTINUITY OF CURRENT

Ampere-Maxwell law is

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 (I + I_D)$$

The sum of the conduction current I and displacement current I_D , $I + I_D$, has the property of continuity. Although, individually each current may not be continuous.

To illustrate the continuity, let us again consider Fig. 4.18. The conduction current through the connecting wire is I , but clearly there is no conduction current between the plates of the capacitor because no charge can be transported across the gap.

For loop C₁, there is no net flux, i.e. $\Phi_E = 0$

$$\text{or } \frac{d\Phi_E}{dt} = 0$$

Therefore,

$$I + I_D = I + \epsilon_0 \frac{d\Phi_E}{dt} = I + 0 = 0 = I \quad (4.37)$$

For loop C₂, there is no conduction current, $I = 0$

Therefore,

$$I + I_D = 0 + \epsilon_0 \frac{d\Phi_E}{dt} = \epsilon_0 \frac{d\Phi_E}{dt} \quad (4.38)$$

If Q is the charge on each plate and A the cross-sectional area of each plate, then the electric field between the plates is

$$E = \frac{Q}{\epsilon_0 A}$$

From Eq. (4.38), we have

$$I + I_D = \epsilon_0 \frac{d\Phi_E}{dt} = \epsilon_0 \frac{d}{dt} \left(\frac{Q}{\epsilon_0 A} \right) = \frac{dQ}{dt} = I \quad (4.39)$$

Thus, we see that the sum of $I + I_D$ is the same for the loop C_1 (i.e., left of plate P) and the loop C_2 proving its continuity, although I and I_D are not individually continuous.

Characteristics of Displacement Current

- (i) Displacement current is a current only in the sense that it produces magnetic field. It is not related with the motion of charges, as usually the case is with the current.
- (ii) Its magnitude is equal to $\epsilon_0 \frac{d\Phi_E}{dt}$, i.e., ϵ_0 time the rate of change of flux.
- (iii) Displacement current serves the purpose of making the total current continuous across the discontinuity in conduction current.

Example 4.10

A parallel plate capacitor with plate cross-sectional area A and separation between the plates d , is

charged by a steady current I . Consider a plane surface of cross-sectional area $\frac{A}{3}$ parallel to the

plates and drawn symmetrically between the plates. Find the displacement current through this area.

Solution. Let, at the time t the charge on the plate be Q . The electric field between the plates is

$$E = \frac{Q}{\epsilon_0 A}$$

The electric flux $\Phi_E = E \times \frac{A}{3} = \frac{Q}{\epsilon_0 A} \times \frac{A}{3} = \frac{Q}{3\epsilon_0}$ through considered area in between the plates.

Therefore, displacement current $I_D = \epsilon_0 \frac{d}{dt} \left(\frac{Q}{3\epsilon_0} \right) = \frac{1}{3}$

$$\frac{dQ}{dt} = \frac{I}{3}$$

Example 4.11

A parallel plate capacitor with circular plates of radius one metre has a capacitance of 2 nF . At $t = 0$, it is connected for charging in series with a resistor $R = 0.5 \text{ M}\Omega$ across a 3 V battery. Calculate the magnetic field at the point P , in between the plates and halfway between the centre and the periphery of the plates, after 10^{-3} sec .

Solution. The problem is depicted in Fig. 4.19.

The time constant of $R-C$ circuit is

$$\tau = CR = 0.5 \times 10^6 \times 2 \times 10^{-9} = 10^{-3} \text{ sec.}$$

The charge on the capacitor at any instant of time t is

$$\begin{aligned} q(t) &= CV [1 - e^{-t/\tau}] \\ &= 2 \times 10^{-9} \times 3 [1 - e^{-t/10^{-3}}] \\ &= 6 \times 10^{-9} (1 - e^{-t/10^{-3}}) \end{aligned}$$

The electric field in between the plates at time t is

$$E = \frac{q(t)}{\epsilon_0 A}$$

As $r = 1 \text{ m}$

$$\therefore A = \pi(1)^2 = \pi \text{ sq. m.}$$

$$\therefore E = \frac{q(t)}{\pi\epsilon_0}$$

Let us now consider a circular loop of radius $\frac{1}{2} \text{ m}$ parallel to the plates passing through P . The

magnetic field B at all points on this loop is along the loop and of the same value.

The electric flux Φ_E through this loop is given by

$$\begin{aligned} \Phi_E &= E \times \text{area of the loop} \\ &= E \times \pi \times \left(\frac{1}{2} \right)^2 = \frac{\pi E}{4} \end{aligned}$$

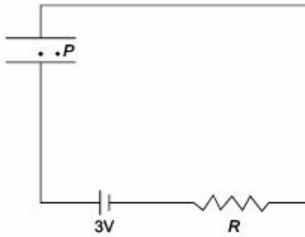


Fig. 4.19

$$= \frac{\pi}{4} \cdot \frac{q(t)}{\pi \epsilon_0} = \frac{q(t)}{4\epsilon_0}$$

The displacement current is

$$\begin{aligned} I_D &= \epsilon_0 \frac{d\Phi_E}{dt} = \epsilon_0 \cdot \frac{d}{dt} \left(\frac{q(t)}{4\epsilon_0} \right) = \frac{1}{4} \frac{dq}{dt} \\ &= \frac{1}{4} \frac{d}{dt} [6 \times 10^{-9} (1 - e^{-t/10^{-3}})] \\ &= 1.5 \times 10^{-9} \times e^{-t/10^{-3}} = 1.5 \times 10^{-6} \times e^{-t/10^{-3}} \end{aligned}$$

\therefore Displacement current at $t = 10^{-3}$, $I_D = 1.5 \times 10^{-6} e^{-1}$

Now, applying Ampere-Maxwell law to this loop, we have

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 (I + I_D)$$

$$B \cdot 2\pi \times \left(\frac{1}{2}\right) = \mu_0 (0 + 1.5 \times 10^{-6} e^{-1})$$

or $B = \frac{\mu_0 \times 1.5 \times 10^{-6} \times e^{-1}}{\pi}$

On putting values, we have

$$\begin{aligned} B &= \frac{4\pi \times 10^{-7} \times 1.5 \times 10^{-6} \times 0.368}{\pi} \\ &= 2.208 \times 10^{-13} \text{ Tesla} \end{aligned}$$

Example 4.12

Fig. 4.20 shows a parallel plate capacitor made of two circular plates, each of radius 10 cm and separated by 5 mm, being charged by a steady current of 0.2 A.

- (a) Calculate the capacitance and rate of change of potential difference between the plates.
- (b) Obtain the displacement current across the plates
- (c) Is Kirchhoff's first rule valid at each plate of the capacitor? Explain.

Solution. (a) The capacitance of a parallel plate capacitor is

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

Here $A = \pi \times (0.1)^2 \text{ m}^2$, $d = 5 \times 10^{-3} \text{ m}$

$$\begin{aligned} \therefore C &= \frac{8.85 \times 10^{-12} \times \pi \times 0.1 \times 0.1}{5 \times 10^{-3}} \text{ F} \\ &= 55.62 \times 10^{-12} \text{ F} \end{aligned}$$

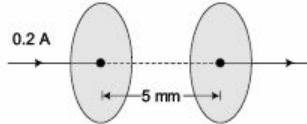


Fig. 4.20

or $C = 55.62 \text{ pF}$

The rate of change of charge C is

$$\frac{dQ}{dt} = C \frac{dV}{dt}$$

or $I = C \frac{dV}{dt}$

\therefore Rate of change of potential difference between the plates

$$\begin{aligned} \frac{dV}{dt} &= \frac{I}{C} = \frac{0.2}{55.62 \times 10^{-12}} \text{ V/s} \\ &= 3.59 \times 10^9 \text{ V/s} \end{aligned}$$

(b) Displacement current

$$\begin{aligned} I_D &= \epsilon_0 \frac{d\Phi_E}{dt} = \epsilon_0 \frac{d}{dt} (EA) = \epsilon_0 A \frac{dE}{dt} \\ &= \epsilon_0 A \frac{d}{dt} \left(\frac{Q}{\epsilon_0 A} \right) = \frac{dQ}{dt} \quad \left(\because E = \frac{Q}{\epsilon_0 A} \right) \end{aligned}$$

$\therefore I_D = I = 0.2 \text{ A}$

(c) Yes, as we see above I (conduction current in the wire) = I_D (displacement current between the plates).

Example 4.13

In Example 4.12 (a) Calculate the magnetic field between the plates at a point (i) on the axis and (ii) 8 cm from the axis, (iii) 15 cm from the axis. (b) At what distance from the axis is the magnetic field due to displacement current greatest? Find the maximum value of the field.

Solution. Let us imagine a circle of radius r between the plates and co axial with them, that is, its centre lies on the axis of the plates and its plane is normal to the axis. By symmetry, the magnetic field \mathbf{B} is tangential to the circle at every point and equal in magnitude over the circle. Therefore, from Ampere's law

$$\oint \mathbf{B} \cdot d\mathbf{l} = 2\pi r B$$

$$\oint \mathbf{B} \cdot d\mathbf{l} = B \cdot 2\pi r = \mu_0 \times \text{current}$$

Here, current is the displacement current enclosed by the circle (of radius r) between the plates.

$$\begin{aligned} \therefore B \cdot 2\pi r &= \mu_0 I_D \cdot \frac{\pi r^2}{\pi R^2} = \mu_0 I_D \frac{r^2}{R^2} \text{ for } r \leq R, \pi r^2 = \text{area of considered circle} \\ &= \mu_0 I_D \text{ for } r \geq R \\ \pi R^2 &= \text{Area of plate} \end{aligned}$$

(a) (i) on the axis $r = 0$

$$B = \frac{\mu_0 I_D}{2\pi} \frac{r}{R^2} = 0$$

(ii) $r = 8 \text{ cm}$

$$\therefore B = \frac{\mu_0 I_D r}{2\pi R^2} = \frac{4\pi \times 10^{-7} \times 0.2 \times 8 \times 10^{-2}}{2\pi \times (0.1)^2} \quad (\because I_D = I = 0.2 \text{ A}) \\ = 3.2 \times 10^{-7} \text{ T}$$

(iii) $r = 16 \text{ cm}$ which is greater than R (10 cm)

$$\therefore B = \frac{\mu_0 I_D}{2\pi r^2} = \frac{4\pi \times 10^{-7} \times 0.2}{2\pi \times 0.16} = 2.5 \times 10^{-7} \text{ T}$$

(b) Maximum value of B is at $r = R$

$$\therefore B_{\max} = \frac{\mu_0 I_D}{2\pi R} = \frac{4\pi \times 10^{-7} \times 0.2}{2\pi \times (0.1)} \text{ T} = 4.0 \times 10^{-7} \text{ T}$$

4.8 MAXWELL'S EQUATIONS

The entire theory of electromagnetic waves is contained in the equations called Maxwell's equations. These four equations are as follows:

$$(i) \oint \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0} \quad (4.40)$$

This is *Gauss's law in electrostatics*. It states that the net electric flux through a closed surface S equals the total charge Q enclosed by the surface divide by ϵ_0 .

$$(ii) \oint \mathbf{B} \cdot d\mathbf{S} = 0 \quad (4.41)$$

This is *Gauss's law in magnetostatics* and states that the net magnetic flux through a closed surface is zero, i.e., the flux entering the surface is equal to the flux leaving it. Alternatively, we can say that law states that *isolated magnetic monopoles do not exist*.

$$(iii) \oint \mathbf{E} \cdot d\mathbf{l} = - \frac{d}{dt} \oint \mathbf{B} \cdot d\mathbf{S} = - \frac{d\Phi_B}{dt} \quad (4.42)$$

This is *Faraday's laws of electromagnetic induction* and states that line integral of electric field around a closed path (or e.m.f.) is equal to the rate of change of magnetic flux through the surface bounded by the closed path.

$$(iv) \oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 I + \mu_0 \epsilon_0 \cdot \frac{d}{dt} \oint \mathbf{E} \cdot d\mathbf{S} \quad (4.43)$$

or $\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 (I + I_D)$

where $I_D = \epsilon_0 \frac{d}{dt} \oint \mathbf{E} \cdot d\mathbf{S} = \epsilon_0 \frac{d\Phi_E}{dt}$

This is *Ampere-Maxwell's law*.

This shows the existence of displacement current and indicates that the magnetic field is produced by both the conduction current and displacement current.

Differential form of Maxwell's equations

The four Maxwell's equations in differential form are given below.

(i) Differential form of Gauss's law is electrostatics

$$\begin{aligned}\operatorname{div} \mathbf{E} &= \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \\ \operatorname{div} \mathbf{D} &= \nabla \cdot \mathbf{D} = \rho\end{aligned}\quad (4.44)$$

where \mathbf{D} ($= \epsilon \mathbf{E}$) is the electric displacement vector $\left(= \frac{Q}{4\pi r^2} \right)$ and ρ is the charge density, i.e., charge per unit volume enclosed by the closed surface.

(ii) Differential form of Gauss's law is magnetostatics

$$\operatorname{div} \mathbf{B} = \nabla \cdot \mathbf{B} = 0 \quad (4.45)$$

(iii) Differential form of Faraday's law of electromagnetic induction.

$$\operatorname{Curl} \mathbf{E} = \nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t} \quad (4.46)$$

(iv) Ampere-Maxwell's law

$$\operatorname{Curl} \mathbf{B} = \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (4.47)$$

or

$$\operatorname{Curl} \mathbf{H} = \nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad (B = \mu_0 H)$$

4.9 WAVE EQUATION FOR PLANE ELECTROMAGNETIC WAVE IN VACUUM

Wave Equation. We have seen in the derivation for the mechanical waves that a function $y(x, t)$, representing the displacement at any point in a mechanical wave travelling along the x -axis must satisfy the following differential equation

$$\frac{\partial^2 y(x, t)}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 y(x, t)}{\partial t^2} \quad (4.48)$$

This equation is called the *wave equation* and v is the speed of propagation of the wave. We shall obtain the corresponding equation for a plain electromagnetic wave in free space or vacuum. First, let us try to understand what we mean by a plane wave.

Plane Wave. As shown in Fig. 4.22, imagine a wavefront of an electromagnetic wave being located, at any instant of time, in a plane perpendicular to the x -axis (parallel to the y - z plane). We assume that the electric field \mathbf{E} has only a y -component and the magnetic field has only a z -component and both fields move together in the $+x$ direction with a speed c . At every point to the left of this plane there are a uniform electric field \mathbf{E} in the $+y$ -direction and a uniform magnetic field in the $+z$ direction. The boundary plane or the wavefront moves to the right in the $+x$ direction with a constant speed c . Thus, the \mathbf{E} and \mathbf{B} fields move to the right into previously field free regions with a definite speed. *A wave like this, in which at any instant, the fields are uniform over any plane perpendicular to the direction of propagation, is called a plane wave.*

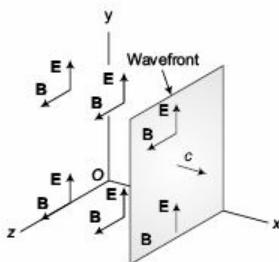


Fig. 4.21 An electromagnetic wavefront. The plane representing the wavefront moves along x-axis with speed c .

In the case shown in Fig. 4.21, the fields are zero for the planes to the right of the wavefront and have the same values on all planes to the left of the wavefront. Thus, a plane wave is the most simple wave to deal with mathematically as complex waves are mathematically complicated.

The followings are the universal features of the electromagnetic waves.

- (i) The wave is transverse, i.e., both \mathbf{E} and \mathbf{B} are perpendicular to the direction of propagation of the wave. The direction of propagation is given by the vector cross product $\mathbf{E} \times \mathbf{B}$. The electric and magnetic fields are also perpendicular to each other.
- (ii) The electromagnetic waves do not require any medium for their propagation unlike mechanical waves in which the particles of the medium oscillate to transmit waves. It is the electric and magnetic fields which 'wave' in an electromagnetic wave.
- (iii) The wave travels in vacuum with a definite and unchanging fields.
- (iv) There is definite ratio between the magnitudes of \mathbf{E} and \mathbf{B} ; $E = cb$

Electromagnetic waves have the property of polarization. A wave in which E is always parallel to certain axis is said to be **linearly polarized** along that axis.

Wave Equation for \mathbf{E} and \mathbf{B} . In vacuum or free space there is no charge or current. Thus, for a region where $\rho = 0$ and $J = 0$, Maxwell's equation becomes

$$\left. \begin{array}{ll} \nabla \cdot \mathbf{E} = 0 & \text{(i)} \\ \nabla \cdot \mathbf{B} = 0 & \text{(ii)} \\ \nabla \times \mathbf{E} = \frac{\partial \mathbf{B}}{\partial t} & \text{(iii)} \\ \nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} & \text{(iv)} \end{array} \right\} \quad (4.49)$$

These equations constitute a set of coupled, first-order, partial differential equations. Taking curl of (iii), we get

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t} \right)$$

Since $\nabla \times (\nabla \times \mathbf{E}) = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ and we see above $\nabla \cdot \mathbf{E} = 0$, therefore, the above expression reduces to

$$\begin{aligned} \nabla (\nabla \cdot \mathbf{E}) &= \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t} \right) \\ \text{or } \nabla^2 \mathbf{E} &= -\frac{\partial}{\partial t} (\nabla \times \mathbf{B}) \\ &= -\frac{\partial}{\partial t} \left(\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) && \text{(using (iv))} \\ \nabla^2 \mathbf{E} &= -\mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} && (4.50) \end{aligned}$$

Similarly, taking curl of (iv) yields

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla \times \left(\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right)$$

Since

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}$$

and as $\nabla \cdot \mathbf{B} = 0$, therefore, the above expression reduces to

$$\begin{aligned} -\nabla^2 \mathbf{B} &= \nabla \times \left(\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \\ &= \mu_0 \epsilon_0 \frac{\partial}{\partial t} (\nabla \times \mathbf{E}) \\ &= -\mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} && \text{(using (iii))} \\ \text{or } \nabla^2 \mathbf{B} &= \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} && (4.51) \end{aligned}$$

Equations. (4.50) and (4.51) are second-order differential equations satisfying the general wave equation given in Eq. (4.48). Thus, E and B behave as a wave while the electromagnetic wave propagates. This implies that the Maxwell's equations supports the propagation of electromagnetic waves through vacuum with a constant speed. Comparing Eqs. (4.50) and (4.48), we get

$$\begin{aligned} \frac{1}{v^2} &= \mu_0 \epsilon_0 \\ \text{or } v &= \frac{1}{\sqrt{\mu_0 \epsilon_0}} && (4.52) \end{aligned}$$

for free space,

$$\mu_0 = 4\pi \times 10^{-7} \text{ N/A}^2$$

and

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

Substituting these values

$$v = \frac{1}{\sqrt{4\pi \times 10^{-7} \text{ N/A}^2 \times 8.85 \times 10^{-12} \text{ C}^2/\text{N.m}^2}}$$

$$\approx 3.00 \times 10^8 \text{ m/s}$$

which is the same as the velocity of light c in free space. This implies that light is an electromagnetic wave, which was a significant inference when this was first calculated although it may appear today a simple conclusion. The result also highlights the importance of Maxwell's modification of Ampere's law, without which the development of electromagnetic theory was not possible.

4.10 SINUSOIDAL ELECTROMAGNETIC WAVES

In a sinusoidal electromagnetic wave, the field vectors E and B at any point in space are sinusoidal function of time and at any instant of time the spatial variation of the fields is also sinusoidal. Some sinusoidal electromagnetic waves are plane waves as they have the property that any instant the fields are uniform over any plane perpendicular to the direction of propagation. Electromagnetic waves produced by an oscillating point charge are an example of sinusoidal waves that are not plane waves. However, if we restrict our attention to a small region of space at sufficiently large distance from the source, even these waves can be approximate to plane waves. This is similar to a small portion of the earth appearing flat of a nearly spherical earth we consider here plane waves for simplicity.

Figure 4.22 shows a linearly polarized sinusoidal electromagnetic wave travelling in the $+x$ direction. It can be noted that the electric and magnetic fields oscillate in phase, i.e., E is maximum where B is maximum and E is zero where B is zero. Also, where E is in $+y$ -direction, B is in $+z$ -direction and where E is in $-y$ -direction, B is in $-z$ -direction. At all points, the vector product $E \times B$ is in the direction in which the wave is propagating ($+x$ -direction). Thus, satisfying the characteristics of electromagnetic waves.

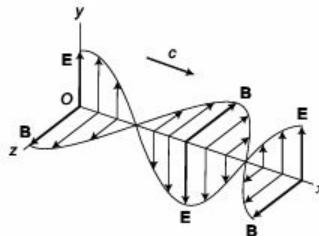


Fig. 4.22 Electric and magnetic fields as functions of x for a linearly polarized sinusoidal plane electromagnetic wave.

Sinusoidal electromagnetic waves can be represented by wave functions, just like the waves on a string. A transverse wave travelling in the $+x$ -direction along a stretched string is given by

$$y(x, t) = A \cos(kx - \omega t) \quad (4.53)$$

where $y(x, t)$ is the transverse displacement from its equilibrium position at time t of a point x on the string. The quantity A is the maximum displacement or *amplitude* of the wave, k is the wave

number $\left(= \frac{2\pi}{\lambda} \right)$ and ω is the angular frequency ($= 2\pi f$, f being the frequency).

Let $E_y(x, t)$ and $B_z(x, t)$ be the instantaneous values of the y -component of \mathbf{E} and z -component of \mathbf{B} , respectively (Fig. 4.23). Let E_{\max} and B_{\max} represent the amplitudes of these fields. Therefore, the wave functions in $+x$ direction are

$$\left. \begin{aligned} E_y(x, t) &= E_{\max} \cos(kx - \omega t) \quad (i) \\ B_z(x, t) &= B_{\max} \cos(kx - \omega t) \quad (ii) \end{aligned} \right\} \quad (4.54)$$

The above expression can also be expressed as

$$\left. \begin{aligned} E(x, t) &= \hat{\mathbf{j}} E_{\max} \cos(kx - \omega t) \quad (i) \\ B(x, t) &= \hat{\mathbf{k}} B_{\max} \cos(kx - \omega t) \quad (ii) \end{aligned} \right\} \quad (4.55)$$

where $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are the unit vectors along y - and z -axes respectively. Eqs. (4.54) and (4.55) show that at any point the sinusoidal oscillations of \mathbf{E} and \mathbf{B} are in phase and the amplitudes are related by

$$E_{\max} = c B_{\max} \quad (4.56)$$

A wave travelling in negative x -direction can be represented as

$$\left. \begin{aligned} E_y(x, t) &= E_{\max} \cos(kx + \omega t) \\ B_z(x, t) &= -B_{\max} \cos(kx + \omega t) \end{aligned} \right\}$$

when \mathbf{E} is in the $+y$ -direction, \mathbf{B} is in $-z$ -direction while when \mathbf{E} is in $-y$ -direction, \mathbf{B} is in $+z$ -direction such that the vector product $\mathbf{E} \times \mathbf{B}$ points in the direction of propagation (Fig. 4.23).

In the exponential from the above mentioned expressions in Eq. (4.54) can be represented as

$$\left. \begin{aligned} E(x, t) &= \hat{\mathbf{j}} E_{\max} e^{i(kx - \omega t)} \quad (i) \\ B(x, t) &= \hat{\mathbf{k}} B_{\max} e^{i(kx - \omega t)} \quad (ii) \end{aligned} \right\} \quad (4.57)$$

which on taking the real part gives the same expressions. We need not always take the $+x$ direction as the direction of propagation. If \mathbf{k} is propagation (or wave) vector pointing in the direction of propagation (whose magnitude is k), then \mathbf{k}, \mathbf{r} is appropriate generalization for kx so that

$$\left. \begin{aligned} \mathbf{E}(x, t) &= \hat{\mathbf{n}} E_{\max} e^{i(\mathbf{k}\mathbf{r} - \omega t)} \quad (i) \\ \mathbf{B}(x, t) &= (\hat{\mathbf{k}} \times \hat{\mathbf{n}}) B_{\max} e^{i(\mathbf{k}\mathbf{r} - \omega t)} = \frac{1}{c} \hat{\mathbf{k}} \times \mathbf{E} \quad (ii) \end{aligned} \right\} \quad (4.58)$$

where $\hat{\mathbf{n}}$ is the polarization vector (the direction of the \mathbf{E} field vector). The actual (real) electric and magnetic field, of a plane sinusoidal electromagnetic wave can be expressed as

$$\left. \begin{aligned} \mathbf{E}(r, t) &= \hat{\mathbf{n}} E_{\max} \cos(\mathbf{k} \cdot \hat{\mathbf{r}} - \omega t) \quad (i) \\ \mathbf{B}(r, t) &= (\hat{\mathbf{k}} \times \hat{\mathbf{n}}) B_{\max} \cos(\mathbf{k} \cdot \hat{\mathbf{r}} - \omega t) \quad (ii) \end{aligned} \right\} \quad (4.59)$$

Example 4.14

A laser emits a sinusoidal electromagnetic wave of wavelength $10.6 \mu\text{m}$ travelling in vacuum along $-x$ -axis. The electric field of maximum magnitude 1.5 M V/m is parallel to the $+z$ -axis. Write expressions for \mathbf{E} and \mathbf{B} as a function of time and position.

Solution. The possible expressions are

$$\mathbf{E}(x, t) = \hat{\mathbf{k}} E_{\max} \cos(kx + \omega t)$$

and

$$\mathbf{B}(x, t) = \hat{\mathbf{j}} B_{\max} \cos(kx + \omega t)$$

Given

$$E_{\max} = 1.5 \text{ M V/m}$$

$$\therefore B_{\max} = \frac{E_{\max}}{c} = \frac{1.5 \times 10^6 \text{ V/m}}{3 \times 10^8 \text{ m/s}} = 5.0 \times 10^{-3} \text{ T}$$

Here,

$$\lambda = 10.6 \mu\text{m} = 10.6 \times 10^{-6} \text{ m}$$

$$\therefore \text{Wave number } k = \frac{2\pi}{\lambda} = \frac{2\pi \text{ rad}}{10.6 \times 10^{-6} \text{ m}} = 5.93 \times 10^5 \text{ rad/m}$$

$$\text{Angular speed } \omega = ck = 3 \times 10^8 \text{ m/s} \times 5.93 \times 10^5 \text{ rad/m} = 1.78 \times 10^{14} \text{ rad/s}$$

Substituting these values, we get

$$\mathbf{E}(x, t) = \hat{\mathbf{k}} (1.5 \times 10^6 \text{ V/m}) \cos [(5.93 \times 10^5 \text{ rad/m})x + (1.78 \times 10^{14} \text{ rad/s})t]$$

$$\mathbf{B}(x, t) = \hat{\mathbf{j}} (1.5 \times 10^{-3} \text{ T}) \cos [(5.93 \times 10^5 \text{ rad/m})x + (1.78 \times 10^{14} \text{ rad/s})t].$$

4.11 ENERGY AND MOMENTUM IN ELECTROMAGNETIC WAVES

Energy is associated with electromagnetic waves, i.e., the electromagnetic waves carry energy. Practical applications of electromagnetic waves such as microwave ovens, radio transmitters, eye surgery using lasers are possible because the electromagnetic waves carry energies. Above all, the sun's radiation carry energy which makes life possible on the earth.

It is known that the energy density associated with an electric field \mathbf{E} is given by

$$u_E = \frac{1}{2} \epsilon_0 E^2$$

where ϵ_0 is the permittivity of free space.

While the energy density associated with a magnetic field \mathbf{B} is given by

$$u_M = \frac{1}{2} \frac{B^2}{\mu_0}$$

where μ_0 is the permeability of the free space.

Therefore, in a region of free space where both electric \mathbf{E} and magnetic \mathbf{B} fields are present, the total energy density is given by

$$u = \frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \frac{B^2}{\mu_0} \quad (4.60)$$

For electromagnetic waves in free space, the magnitudes of \mathbf{E} and \mathbf{B} are related by

$$B = \frac{E}{c} = \sqrt{\mu_0 \epsilon_0} E \quad (4.61)$$

Therefore, the total energy density can also be expressed as

$$u = \frac{1}{2} \epsilon_0 E^2 + \frac{1}{2\mu_0} (\sqrt{\mu_0 \epsilon_0} E)^2 = \frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \epsilon_0 E^2 = \epsilon_0 E^2 \quad (4.62)$$

This shows that in free space the energy density associated with the **E** field in the electromagnetic wave is equal to the energy density associated with the **B** field. In general, the electric field magnitude **E** is a function of position and time, as for sinusoidal electromagnetic waves, and therefore the energy density **u** of an electromagnetic wave also depends on position and time.

Energy flow and Poynting Vector. Electromagnetic waves transport energy as they propagate from one region to another. This transfer of energy can be described in terms of *energy transferred per unit time per unit cross-sectional area* or *power per unit area* for an area perpendicular to the direction of propagation of the wave.

Let us consider a stationary plane perpendicular to the x-axis (Fig. 4.23), that coincides with the wavefront at a certain time. In the time interval dt after this, the wavefront moves a distance $dx = c dt$ to the right of the plane. Considering an area **A** on this stationary plane, energy in the space to the right of this area must have passed through the area to reach the new location. The volume dV of this region is area **A** times the length $c dt$, and the energy dU in this region is the energy density **u** times the volume

$$dV = u dV = (\epsilon_0 E^2) (A c dt)$$

This energy passes through this area **A** in time dt . Therefore, the energy flow per unit time per unit area, denoted by **S** is, given by

$$S = \frac{1}{A} \frac{dU}{dt} = \epsilon_0 c E^2 \quad (4.63)$$

Alternatively, substituting $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$, we have

$$S = \frac{\epsilon_0}{\sqrt{\mu_0 \epsilon_0}} \cdot E^2 = \sqrt{\frac{\epsilon_0}{\mu_0}} \cdot E^2 = \frac{EB}{\mu_0} \quad (\because E = c B) \quad (4.64)$$

The units of **S** are energy per unit time per unit area or power per unit area. The SI unit of **S** is 1 J/s. m² or 1 W/m².

We define a vector quantity that gives both the magnitude and direction of the energy flow rate,

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \quad (4.65)$$

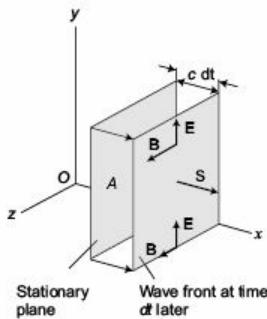


Fig. 4.23 A wave front at a time dt after it passes through the stationary plane with a speed c .

The vector \mathbf{S} is called the *Poynting vector*, named after British physicist John Poynting. The magnitude of \mathbf{S} is equal to $\frac{EB}{\mu_0}$ while its direction is given by the vector product $\mathbf{E} \times \mathbf{B}$, i.e., perpendicular to both \mathbf{E} and \mathbf{B} , in the direction of propagation of the wave. Thus, the *Poynting vector gives the energy flow per unit area per unit time through a cross-sectional area perpendicular to the direction of propagation*.

The total energy flow per unit time (= Power P) out of any closed surface is the integral of \mathbf{S} over the surface S ,

$$P = \oint \mathbf{S} \cdot d\mathbf{A} \quad (4.66)$$

For the sinusoidal waves and other complicated waves, the electric and magnetic fields at any point vary with time so the Poynting vector at any point is also a function of time. Because of the frequencies of the typical electromagnetic waves are very high, the time variation of the Poynting vector is so rapid that is most appropriate to deal with the average value. *The magnitude of the average value of \mathbf{S} at a point is called the intensity of the radiation at the point.* Let $E_y(x, t)$ and $B_z(x, t)$ represent the instantaneous values of y -component of \mathbf{E} and z -component of \mathbf{B} respectively of a sinusoidal electromagnetic wave.

Thus,

$$\left. \begin{aligned} \mathbf{E}(x, t) &= \hat{\mathbf{j}} E_{\max} \cos(kx - \omega t) \quad (\text{i}) \\ \mathbf{B}(x, t) &= \hat{\mathbf{k}} B_{\max} \cos(kx - \omega t) \quad (\text{ii}) \end{aligned} \right\} \quad (4.67)$$

where E_{\max} and B_{\max} are the maximum amplitudes of the electric and magnetic fields respectively.

It is important to remember that $\hat{\mathbf{k}}$ is a unit vector in the z -direction while k is the wave number. The intensity of the sinusoidal wave is therefore given by

$$\begin{aligned}\mathbf{S}(x, t) &= \frac{1}{\mu_0} \mathbf{E}(x, t) \times \mathbf{B}(x, t) \\ &= \frac{1}{\mu_0} [\hat{\mathbf{j}} E_{\max} \cos(kx - \omega t)] \times [\hat{\mathbf{k}} B_{\max} \cos(kx - \omega t)]\end{aligned}\quad (4.68)$$

As $\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}$ and $\cos^2(kx - \omega t)$ is always positive, so $\mathbf{S}(x, t)$ always points in the positive x -direction, the direction of wave propagation. The x -component of the Poynting vector is

$$\begin{aligned}\mathbf{S}_x(x, t) &= \frac{E_{\max} B_{\max}}{\mu_0} \cos^2(kx - \omega t) \\ &= \frac{E_{\max} B_{\max}}{2\mu_0} [1 + \cos^2(kx - \omega t)]\end{aligned}\quad (4.69)$$

The time average value of $\cos^2(kx - \omega t)$ is zero because at any point it is positive during on half-cycle and negative during the other half. So the average value of the Poynting vector over a full cycle is $\mathbf{S}_{av} = \hat{\mathbf{i}} S_{av}$, where

$$S_{av} = \frac{E_{\max} B_{\max}}{2\mu_0} \quad (4.70)$$

That is, the magnitude of the average value of \mathbf{S} for a sinusoidal wave is half times the maximum value. As, $E_{\max} = c B_{\max}$, the intensity can be expressed as

$$I = S_{av} = \frac{E_{\max} B_{\max}}{2\mu_0} = \frac{E^2_{\max}}{2\mu_0 c} = \frac{1}{2} \sqrt{\frac{\epsilon_0}{\mu_0}} E^2_{\max} = \frac{1}{2} \epsilon_0 c E^2_{\max} \quad (4.71)$$

You can easily show that for a wave travelling in the $-x$ -direction, the Poynting vector is in the $-x$ -direction at every point, its magnitude remains the same as for a wave travelling in the $+x$ direction.

We have derived the above relations for free space. However, in any dielectric medium just replace ϵ_0 by ϵ (the permittivity of the dielectric) and μ_0 by μ (the permeability of the dielectric) and c by v (velocity in the medium) and you will get the expressions for the dielectric medium. Even the energy densities in the \mathbf{E} and \mathbf{B} fields are equal in a dielectric medium.

Momentum flow and Radiation Pressure

We have seen that electromagnetic waves carry energy. It can also be shown that electromagnetic waves carry momentum p , with a corresponding momentum density (momentum dp per volume dV) of magnitude

$$\frac{dp}{dV} = \frac{EB}{\mu_0 c^2} = \frac{S}{c^2} \quad (4.72)$$

Unlike in the mechanical wave where the momentum is associated with the mass of a moving particle, in case of electromagnetic waves the momentum is associated with the field.

The volume dV occupied by the electromagnetic wave that passes through an area A in time dt is $dV = c Adt$. Substituting this in Eq. (4.72), we find that the *momentum flow rate per unit area* is

$$\frac{1}{A} \frac{dp}{dt} = \frac{S}{c} = \frac{EB}{\mu_0 c} \quad (4.73)$$

This is the *momentum transferred per unit surface area per unit time*. We can obtain the *average rate of momentum transfer per unit area* by replacing S in Eq. (4.73) by $S_{av} = I$.

This momentum transfer is responsible for the phenomenon of *radiation pressure*. When an electromagnetic wave is completely absorbed by a surface, the wave's momentum is also transferred to the surface. For simplicity, we consider the surface perpendicular to the propagation direction.

We know, that the rate $\frac{dp}{dt}$ at which momentum is transferred to the absorbing surface equals the force on the surface.

The *average force per unit area due to the wave or radiation pressure* p_{rad} is the average value of $\frac{dp}{dt}$ divided by the absorbing area A (we have used subscript rad to distinguish radiation pressure from the momentum p). Thus, the radiation pressure is given by

$$P_{rad} = \frac{S_{av}}{c} = \frac{I}{c} \quad (4.74)$$

In case the wave is totally reflected the momentum change is twice p_{rad} given by

$$P_{rad} = \frac{2S_{av}}{c} = \frac{2I}{c} \quad (4.75)$$

For example, the value of S_{av} (or I) for direct sunlight before it passes through the earth's atmosphere is approximately 1.4 kW/m^2 . The corresponding radiation pressure on a completely absorbing surface is

$$P_{rad} = \frac{S_{av}}{c} = \frac{1.4 \times 10^3 \text{ W/m}^2}{3.0 \times 10^8 \text{ m/s}} = 4.7 \times 10^{-6} \text{ Pa}$$

For a totally reflecting surface $p_{rad} = 2 \times 4.7 \times 10^{-6} \text{ Pa}$. These are very small pressures of the order of 10^{-10} atmosphere but can be measured using a highly sensitive instrument.

The radiation pressure inside the stars is very high. Inside the stars, which are highly massive and luminous than the sun, the radiation pressure is so high that it substantially augments the gas pressure within the star and helps the star from collapsing under its own gravity. Also, from time to time some stars undergo an enormous increase in its brightness. When this happens, radiation pressure at the star's surface becomes so strong that parts of its outer layers are ejected into space. Some of this ejected material appears as glowing blobs surrounding the star.

4.12 POYNTING'S THEOREM: ENERGY IN ELECTROMAGNETIC WAVES

Let us consider an electromagnetic wave propagating along $+x$ axis, with the electric and magnetic fields \mathbf{E} and \mathbf{B} confined in the transmission planes along y and z -axis respectively. The direction of propagation is also given by the propagation vector $(\mathbf{E} \times \mathbf{B})$. Taking the divergence of the Poynting vector in free space

$$\begin{aligned} \nabla \cdot (\mathbf{E} \times \mathbf{B}) &= \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B}) \\ &= -\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mu_0 \epsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \quad (\text{using Eq. (4.49)}) \end{aligned}$$

As $\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} = \frac{1}{2} \frac{\partial}{\partial t} (B^2)$ and $\mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{2} \frac{\partial (E^2)}{\partial t}$, we have

$$\begin{aligned}\nabla \cdot (\mathbf{E} \times \mathbf{B}) &= -\frac{1}{2} \frac{\partial(B^2)}{\partial t} - \frac{1}{2} \mu_0 \epsilon_0 \frac{\partial(E^2)}{\partial t} \\ &= -\frac{\partial}{\partial t} \left(\frac{1}{2} \mu_0 \epsilon_0 E^2 + \frac{1}{2} \frac{B^2}{\mu_0} \right) \\ \text{or } \frac{1}{\mu_0} \nabla \cdot (\mathbf{E} \times \mathbf{B}) &= -\frac{\partial}{\partial t} \left(\frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \frac{B^2}{\mu_0} \right)\end{aligned}\quad (4.76)$$

Considering a surface S bounding a volume V and integrating the above relation over the volume V , we get

$$\frac{1}{\mu_0} \int_V \nabla \cdot (\mathbf{E} \times \mathbf{B}) = -\frac{\partial}{\partial t} \left[\int_V \left(\frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \frac{B^2}{\mu_0} \right) dV \right]$$

Applying the divergence theorem on the L.H.S., we get

$$\frac{1}{\mu_0} \int_S (\mathbf{E} \times \mathbf{B}) \cdot d\mathbf{a} = -\frac{\partial}{\partial t} \int_V \left(\frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \frac{B^2}{\mu_0} \right) dV \quad (4.77)$$

The term on the R.H.S. is called the Poynting vector and represents the *energy transported by the electromagnetic fields per unit time per unit area*. Thus, the Poynting vector \mathbf{S} given by

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \quad (4.78)$$

gives the magnitude and direction of energy flow rate. Since \mathbf{E} and \mathbf{B} are perpendicular, the magnitude of \mathbf{S} is $S = \frac{EB}{\mu_0}$. This gives the energy flow per unit area per unit time through a cross-sectional area perpendicular to the propagation direction. The instantaneous energy density associated with the electric field \mathbf{E} is given by

$$u_E = \frac{1}{2} \epsilon_0 E^2 \quad (4.79)$$

While that with the magnetic field \mathbf{B} is given by

$$u_B = \frac{1}{2} \frac{B^2}{\mu_0} \quad (4.80)$$

The sum of the two terms on the R.H.S. $\left(\frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \frac{B^2}{\mu_0} \right)$ represents the energy associated with the electromagnetic wave.

$$\begin{aligned}\text{Also, as } E &= c B, u_E = \frac{1}{2} \epsilon_0 E^2 = \frac{1}{2} \epsilon_0 (Bc)^2 \\ &= \frac{1}{2} \epsilon_0 B^2 \cdot c^2 = \frac{1}{2} \epsilon_0 \frac{B^2}{\mu_0 \epsilon_0}\end{aligned}$$

$$= \frac{1}{2} \frac{B^2}{\mu_0}.$$

That is, the instantaneous energy density with electric field is equal to the instantaneous energy density associated with the magnetic field.

The expression given by Eq. (4.77) represents the *Poynting theorem* according to which the energy flow per unit time (= power P) out of any closed surface is the integral of S over the surface

$$P = \oint S \cdot d\mathbf{a} \quad (4.81)$$

Example 4.15

In an electromagnetic wave, the electric field of amplitude 6.2 V/m oscillates with a frequency of 2.4×10^{10} Hz. (a) Calculate the wavelength of the wave, and (b) what is the energy density of the electromagnetic field of the wave?

Solution. (a) Here, $v = 2.4 \times 10^{10}$ Hz

we know, $c = v\lambda$

$$\therefore \text{wavelength} \quad \lambda = \frac{c}{v} = \frac{3.0 \times 10^8 \text{ m/s}}{2.4 \times 10^{10} \text{ m}} = 1.25 \times 10^{-2} \text{ m}$$

(b) Energy density $u_{\text{em}} = \epsilon_0 E_{\text{max}}^2$

Here, $E_{\text{max}} = 6.2 \text{ V/m}$ and $\epsilon_0 = 8.85 \times 10^{-12} \text{ C}^2/\text{N m}^2$

$$\therefore u_{\text{em}} = 8.85 \times 10^{-12} \times 6.2 \times 6.2 \\ = 3.4 \times 10^{-10} \text{ J/m}^3$$

Example 4.16

Calculate the amplitude of the electric field in a parallel beam of light of intensity 3.0 W m^{-2} .

Solution. We know, intensity $I = \frac{1}{2} \epsilon_0 c E_{\text{max}}^2$

$$\therefore \text{Amplitude of electric field} \quad E_{\text{max}} = \sqrt{\frac{2I}{\epsilon_0 c}} \\ = \sqrt{\frac{2 \times 3.0}{8.85 \times 10^{-12} \times 3 \times 10^8}} \text{ N/C} \\ = 47.53 \text{ N/C}$$

Example 4.17

For an electromagnetic wave $E = 100 \text{ V/m} = 100 \text{ N/C}$. Find the value of B, the energy density and the rate of flow of energy per unit area.

Solution. $B = \frac{E}{c} = \frac{100 \text{ Vm}^{-1}}{3 \times 10^8 \text{ ms}^{-1}} = 3.33 \times 10^{-7} \text{ T}$

Energy density $u_{\text{em}} = \epsilon_0 E^2 = 8.85 \times 10^{-12} \text{ C}^2/\text{Nm}^2 \times 100 \text{ N/C} \times 100 \text{ N/C}$
 $= 8.85 \times 10^{-8} \text{ Jm}^{-3}$ (or Nm^{-2})

$$\begin{aligned}\text{Rate of flow of energy per unit area} &= \text{Magnitude of Poynting vector } S = \frac{EB}{\mu_0} \\ &= \frac{100 \text{ V/m} \times 3.33 \times 10^{-7} \text{ T}}{4\pi \times 10^{-7} \text{ T.m/A}} = 2.65 \text{ W/m}^2 \text{ (or A/m}^2)\end{aligned}$$

Example 4.18

A satellite orbiting the earth has solar energy collecting panels with total surface area of 4 m^2 . If the sun's radiation is incident normally and absorbed completely, calculate the average solar power and the average force associated with radiation pressure. (Take $I = 1.4 \times 10^3 \text{ Wm}^{-2}$)

Solution. Here, $A = 4.0 \text{ m}^2$ and $I = 1.4 \times 10^3 \text{ Wm}^{-2}$

$$\begin{aligned}\text{Power} &= IA = 1.4 \times 10^3 \text{ Wm}^{-2} \times 4.0 \text{ m}^2 \\ &= 5.6 \times 10^3 \text{ W} = 5.6 \text{ kW}\end{aligned}$$

$$\text{Radiation pressure} = p_{\text{rad}} = \frac{I}{c} = \frac{1.4 \times 10^3 \text{ Wm}^{-2}}{3.0 \times 10^8 \text{ m/s}} = 4.7 \times 10^{-6} \text{ N/m}^2$$

$$\begin{aligned}\text{Force} &= p_{\text{rad}} A \\ &= 4.7 \times 10^{-6} \text{ Nm}^{-2} \times 4.0 \text{ m} \\ &= 1.9 \times 10^{-5} \text{ N}\end{aligned}$$

Example 4.19

A radio station radiates a sinusoidal wave with an average total power of 50 kW. Assuming that the transmitter radiates equally in all directions above the ground, find the amplitudes of E_{max} and B_{max} detected by a satellite at a distance 100 km from the radio station transmitter antenna.

Solution. We can take the covered area by the transmitter to be a hemisphere of radius 100 km centred on the transmitter.

$$\therefore \text{Surface area of hemisphere} = 2\pi R^2 = 2\pi \times (100 \times 10^3 \text{ m})^2 = 6.28 \times 10^{10} \text{ m}^2$$

$$\text{Intensity } I = \frac{\text{Power}}{\text{Area}} = \frac{P}{2\pi R^2} = \frac{50 \times 10^3 \text{ W}}{6.28 \times 10^{10} \text{ m}^2} = 7.96 \times 10^{-7} \text{ Wm}^{-2}$$

$$\text{As } I = S_{\text{av}} = \frac{E_{\text{max}}^2}{2\mu_0 c}$$

$$\begin{aligned}\therefore E_{\text{max}} &= \sqrt{2I\mu_0 c} \\ &= \sqrt{2 \times 7.96 \times 10^{-7} \text{ Wm}^{-2} \times 4\pi \times 10^{-7} \frac{\text{Tm}}{A} \times 3 \times 10^8 \text{ m/s}} \\ &= 2.45 \times 10^{-2} \text{ Vm}^{-1}\end{aligned}$$

$$B_{\text{max}} = \frac{E_{\text{max}}}{c} = \frac{2.45 \times 10^{-2} \text{ Vm}^{-1}}{3 \times 10^8 \text{ m/s}} = 8.17 \times 10^{-11} \text{ T}$$

4.13 ELECTROMAGNETIC WAVES IN A DIELECTRIC MEDIUM

So far, we have discussed the propagation of electromagnetic waves through free space. Now let us consider the propagation of electromagnetic waves through a homogeneous and isotropic *non-conducting* or *dielectric* material having permittivity ϵ and permeability μ . All the Maxwell's equations remain the same except ϵ_0 and μ_0 are to be replaced by ϵ and μ . For such a medium

$$\mathbf{B} = \mu H \text{ and } \mathbf{D} = \epsilon \mathbf{E}$$

Also, the electric current density for a dielectric is zero, i.e.,

$$\mathbf{J} = \sigma \mathbf{E} = 0 \text{ and } \rho = 0$$

The Maxwell's equation in this case become

$$\left. \begin{array}{l} \operatorname{div} \mathbf{E} = \nabla \cdot \mathbf{E} = 0 \\ \operatorname{div} \mathbf{B} = \nabla \cdot \mathbf{B} = 0 \\ \operatorname{curl} \mathbf{E} = \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\ \operatorname{curl} \mathbf{B} = \nabla \times \mathbf{B} = \mu \epsilon \frac{\partial \mathbf{E}}{\partial t} \end{array} \right\} \quad (4.82)$$

Differentiating Eq. (4.82) (iv), we get

$$\frac{\partial}{\partial t} (\nabla \times \mathbf{B}) = \frac{\partial}{\partial t} \left(\mu \epsilon \frac{\partial \mathbf{E}}{\partial t} \right)$$

$$\text{or } \nabla \times \frac{\partial \mathbf{B}}{\partial t} = \mu \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}$$

Replacing $\frac{\partial \mathbf{B}}{\partial t}$ by $-\nabla \times \mathbf{E}$, we get

$$\begin{aligned} \nabla \times (-\nabla \times \mathbf{E}) &= \mu \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \\ \text{or } -\nabla \times (\nabla \times \mathbf{E}) &= \mu \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \end{aligned} \quad (4.83)$$

Using the identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ and

$\nabla \cdot \mathbf{E} = 0$, we have

$$\nabla^2 \mathbf{E} = \mu \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (4.84)$$

Similarly, obtain

$$\nabla^2 \mathbf{B} = \mu \epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2} \quad (4.85)$$

These are wave equations for the electric and magnetic field vectors respectively. Comparing with the standard wave equation (Eq. 4.48) the velocity of wave propagation is given by

$$v = \frac{1}{\sqrt{\mu\epsilon}} \quad (4.86)$$

If ϵ_r is the relative permittivity or dielectric constant of the dielectric so that $\epsilon = \epsilon_r \epsilon_0$ and μ_r is the relative permeability of the dielectric so that $\mu = \mu_r \mu_0$, then from Eq. (4.86), the wave speed in the dielectric is also given by

$$\begin{aligned} v &= \frac{1}{\sqrt{\mu\epsilon}} = \frac{1}{\sqrt{\mu_r\mu_0\epsilon_r\epsilon_0}} \\ &= \frac{1}{\sqrt{\mu_0\epsilon_0}} \cdot \frac{1}{\sqrt{\mu_r\epsilon_r}} = \frac{c}{\sqrt{\mu_r\epsilon_r}} \end{aligned} \quad (4.87)$$

For most dielectrics μ_r is very nearly equal to unity (except for insulating ferromagnetic materials). In such cases,

$$v = \frac{c}{\sqrt{\epsilon_r}} \quad (4.88)$$

Because the relative permittivity or dielectric constant is always greater than unity, the speed v of electromagnetic waves in a dielectric is always less than the c , the speed in vacuum by a factor

$$\frac{1}{\sqrt{\epsilon_r}}$$

The ratio of the speed c in vacuum to the speed v in a material is called the *refractive index* n of the material. Therefore,

$$\frac{c}{v} = n = \sqrt{\mu_r\epsilon_r} \cong \sqrt{\epsilon_r} \quad (\text{when } \mu_r \cong 1) \quad (4.89)$$

The dielectric constant ϵ_r of medium is dependent on the frequency of the field. For example, for dc fields, the dielectric constant of water is 80.4 while in the frequency range of visible light it is 1.8 only. It is because in highly oscillating fields the dipoles fail to orient with the oscillating field.

Example 4.20

Light from a sodium lamp of frequency 5.09×10^{14} Hz is used to view a diamond. If for the diamond $\epsilon_r = 5.84$ and $\mu_r = 1.00$ find the speed and wavelength of light in diamond.

Solution. Wavelength in vacuum

$$\lambda_{\text{vac}} = \frac{c}{v} = \frac{3 \times 10^8 \text{ m/s}}{5.09 \times 10^{14} \text{ Hz}} = 5.89 \times 10^{-7} \text{ m} = 589 \text{ nm}$$

$$\text{Speed of light in vacuum} \quad v = \frac{c}{\sqrt{\mu_r\epsilon_r}} = \frac{3 \times 10^8 \text{ m/s}}{\sqrt{5.84}} = 1.24 \times 10^8 \text{ m/s}$$

wavelength in diamond

$$\lambda_d = \frac{v}{n} = \frac{1.24 \times 10^8 \text{ m/s}}{5.09 \times 10^{14} \text{ s}^{-1}} = 2.44 \times 10^{-7} \text{ m} = 244 \text{ nm}$$

Example 4.21

A FM radio channel broadcasts its signal at 98.3 MHz. What is the wavelength of the signal in free space. This signal passes through an insulating ferrite with $\epsilon_r = 10$ and $\mu_r = 1000$. What is speed and wavelength of the signal in ferrite?

$$\text{Solution. Wavelength in free space } \lambda_{\text{vac}} = \frac{c}{v} = \frac{3 \times 10^8 \text{ m/s}}{98.3 \times 10^6 \text{ Hz}} = 3.05 \text{ m}$$

$$\text{Speed in ferrite } v = \frac{c}{\sqrt{\mu_r \epsilon_r}} = \frac{3 \times 10^8 \text{ m/s}}{\sqrt{1000 \times 10}} = 3 \times 10^6 \text{ m/s}$$

$$\text{Wavelength in ferrite } \lambda_{\text{ferrite}} = \frac{v}{c} = \frac{3 \times 10^6 \text{ m/s}}{98.3 \times 10^6 \text{ Hz}} = 3.05 \text{ cm}$$

4.14 ELECTROMAGNETIC WAVES IN A CONDUCTING MEDIUM

Let us consider the propagation of electromagnetic waves through a homogeneous conducting medium of permittivity ϵ , permeability μ , conductivity σ and free charge density ρ_f . According to Ohm's law, the (free) current density \mathbf{J}_f in a conductor is proportional to the electric field.

$$\mathbf{J}_f = \sigma \mathbf{E} \quad (4.90)$$

In this case, Maxwell's equations take the form

$$\left. \begin{array}{l} \nabla \cdot \mathbf{E} = \frac{\rho_f}{\epsilon} \\ \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} = \mu \sigma \mathbf{E} + \mu \epsilon \frac{\partial \mathbf{E}}{\partial t} \end{array} \right\} \quad (4.91)$$

The continuity equation for free charge is given by

$$\nabla \cdot \mathbf{J}_f = -\frac{\partial \rho_f}{\partial t} \quad (4.92)$$

Using Ohm's law $J_f = \sigma E$ and Gauss' law $\nabla \cdot \mathbf{E} = \frac{\rho_f}{\epsilon}$, in the continuity equation, we have

$$\frac{\partial \rho_f}{\partial t} = -\sigma (\nabla \cdot \mathbf{E}) = -\frac{\sigma}{\epsilon} \rho_f$$

The above equation can be integrated and using $\rho_f = \rho_f(0)$ at $t = 0$ obtain

$$\rho_f(t) = \rho_f(0) e^{-\frac{\sigma t}{\epsilon}} \quad (4.93)$$

That is, any initial free charge dissipates in a characteristic time $\tau = \frac{\epsilon}{\sigma}$. This is also manifested by the well known fact that any free charge put on a conductor flow out to the edges. The time constant τ , in a sense, is a measure that how good a conductor is. For a 'perfect' conductor,

conductivity $\sigma = \infty$ and therefore time constant $\tau = 0$. For a 'good' conductor, τ is much less than the other relevant times such as time between collisions, for oscillatory systems, $\tau \ll \frac{1}{\omega}$. For a 'poor' conductor, τ is greater than the characteristic time ($\tau \gg \frac{1}{\omega}$). Thus, in conductors any free charge quickly disappear and the interior of the conductor may be regarded as free of free charges, i.e., $\rho_f = 0$. In other words, there is no sustainable free charge in the conductor. Therefore, the Maxwell's equations becomes

$$\left. \begin{array}{l} \nabla \cdot \mathbf{E} = 0 \\ \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} = -\frac{\sigma \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} = \mu\sigma \mathbf{E} + \mu\epsilon \frac{\partial \mathbf{E}}{\partial t} \end{array} \right\} \quad (4.94)$$

Thus, the only difference, in comparison to the Maxwell's equations for a dielectric medium, is the presence of $\mu\sigma\mathbf{E}$ in the last equation.

Taking curl of (iii) in Eq. (4.94), we get

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \cdot (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = \nabla \times \left(-\frac{\sigma \mathbf{B}}{\partial t} \right)$$

As $\nabla \cdot \mathbf{E} = 0$, we have

$$\nabla^2 \mathbf{E} = \frac{\sigma}{\partial t} (\nabla \times \mathbf{B}) = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{E}}{\partial t} \quad (4.95)$$

Now taking curl of (iv) in Eq. (4.94), we get

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla \cdot (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = \nabla \times \left(\mu\sigma \mathbf{E} + \mu\epsilon \frac{\partial \mathbf{E}}{\partial t} \right)$$

As $\nabla \cdot \mathbf{B} = 0$, we have

$$-\nabla^2 \mathbf{B} = \mu\sigma (\nabla \times \mathbf{E}) + \mu\epsilon \frac{\partial \mathbf{E}}{\partial t} (\nabla \times \mathbf{E})$$

or $\nabla^2 \mathbf{B} = \mu\sigma \frac{\partial \mathbf{B}}{\partial t} + \mu\epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2}$

or $\nabla^2 \mathbf{B} = \mu\epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{B}}{\partial t}$ (4.96)

These are the wave equations for the electric and magnetic field vectors. The solutions of these equations are of the form

$$\mathbf{E}(x, t) = \mathbf{E}_{\max} e^{i(k'x - \omega t)} \quad (4.97)$$

and $\mathbf{B}(x, t) = \mathbf{B}_{\max} e^{i(k'x - \omega t)}$ (4.98)

but now the wave number k' is a complex quantity given by

$$k'^2 = \mu\epsilon\omega^2 + j\mu\sigma\omega \quad (4.99)$$

which can be verified by putting in the wave equation. The square root of the above equation yields

$$k' = k + is \quad (4.100)$$

where,

$$k = \omega \sqrt{\frac{\epsilon\mu}{2}} \left[\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega} \right)^2} - 1 \right]^{\frac{1}{2}} \quad (4.101)$$

and

$$s = \omega \sqrt{\frac{\epsilon\mu}{2}} \left[\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega} \right)^2} - 1 \right]^{\frac{1}{2}} \quad (4.102)$$

The imaginary part of k' results in the attenuation of the wave, i.e., the amplitude decreases with increasing x . Thus, the two fields can be expressed as

$$\mathbf{E}(x, t) = \mathbf{E}_{\max} e^{-sx} e^{i(kx - \omega t)} \quad (4.103)$$

$$\mathbf{B}(x, t) = \mathbf{B}_{\max} e^{-sx} e^{i(kx - \omega t)} \quad (4.104)$$

The above expressions show that the field vectors are damped in a conducting medium (Fig. 4.24). The distance (along the direction of propagation) where the amplitude of the field vector reduces to $\frac{1}{e}$ is called the 'skin depth'. That is,

$$d = \frac{1}{s} \quad (4.105)$$

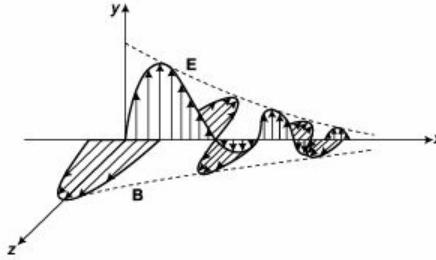


Fig. 4.24 Damped electromagnetic wave.

Thus, the skin depth is a measure of how far the wave penetrate into the conductor. The real part of k' can be used to find the wavelength λ , the wave speed v and the refractive index n using the following usual expressions.

$$\lambda = \frac{2\pi}{k}$$

$$v = \frac{\omega}{k}$$

and

$$n = \frac{ck}{\omega}$$

where k is given by Eq. (4.101).

The attenuated plane waves given by Eqs. (4.103) and (4.104) satisfy the modified wave equation (Eqs. (4.95) and (4.96)) for any E_{\max} and B_{\max} . But Maxwell's equations (Eq. 4.94) impose further constraints, which can be used to determine relative amplitudes, phases and polarizations of \mathbf{E} and \mathbf{B} . Eqs. (4.94(i)) and (4.94(ii)) rules out any x components, i.e., the fields are transverse. The electric field along y -axis may be represented as

$$\mathbf{E}(x, t) = \hat{\mathbf{j}} E_{\max} e^{-ixx} e^{i(kx - \omega t)} \quad (4.106)$$

and the magnetic field

$$\mathbf{B}(x, t) = \hat{\mathbf{k}} \frac{k'}{\omega} E_{\max} e^{-ixx} e^{i(kx - \omega t)} \quad (4.107)$$

Like any complex number, k' can be expressed in terms of its modulus (k) and phase ϕ as

$$k' = K e^{i\phi} \quad (4.108)$$

where

$$K = |k'| = \sqrt{k^2 + s^2} = \omega \sqrt{\epsilon \mu} \sqrt{1 + \left(\frac{\sigma}{\epsilon \omega}\right)^2} \quad (4.109)$$

and

$$\phi = \tan^{-1} \left(\frac{s}{k} \right) \quad (4.110)$$

According to Eqs. (4.106) and (4.107), the complex amplitudes $E = E_{\max} e^{i\delta_E}$ and $B = B_{\max} e^{i\delta_B}$ are related by

$$B_{\max} e^{i\delta_B} = \frac{K e^{i\phi}}{\omega} E_{\max} e^{i\delta_E} \quad (4.111)$$

Obviously, the electric and magnetic fields are no longer in phase and the phase difference is given by

$$\delta_B - \delta_E = \phi \quad (4.112)$$

Thus, the magnetic field lags behind the electric field. From Eq. (4.107), it is evident that the real amplitudes of \mathbf{E} and \mathbf{B} are related by

$$\frac{B_0}{E_0} = \frac{k}{\omega} = \sqrt{\epsilon \mu} \sqrt{1 + \left(\frac{\sigma}{\epsilon \omega}\right)^2} \quad (4.113)$$

Thus, the real electric and magnetic fields can be represented as

$$\mathbf{E}(x, t) = \hat{\mathbf{j}} E_{\max} e^{-ixx} \cos(kx - \omega t)$$

$$\mathbf{B}(x, t) = \hat{\mathbf{k}} E_{\max} e^{-ixx} \cos(kx - \omega t + \phi)$$

These fields are shown in Fig. 4.25.

4.15 THE ELECTROMAGNETIC SPECTRUM

Electromagnetic waves cover a wide range of wavelength or frequency. The electromagnetic spectrum includes radio and TV signals, visible light, infrared and ultraviolet radiation, X-rays and gamma rays. These broad portions are depicted in Fig. 4.25, showing the approximate wavelength and

frequency ranges for each portion. They all have a common property that they propagate with a speed of 3×10^8 m/s, though they differ drastically in the way they are produced and used.

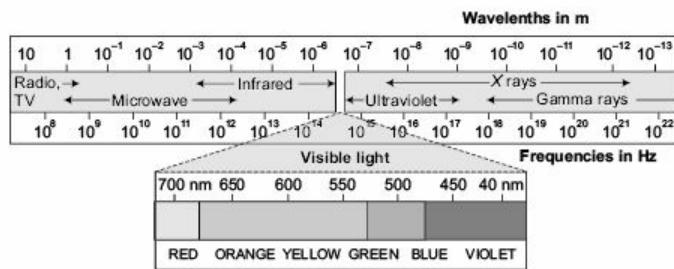


Fig. 4.25 The electromagnetic spectrum (the boundaries between bands are somewhat arbitrary).

Only a small portion of this broad spectrum can be detected by our eyes. We call this small portion *visible light*. Its wavelengths range from about 400 to 700 nm ($1 \text{ nm} = 10^{-9} \text{ m}$), with corresponding frequencies varying from 750 to 430 THz (or 7.5 to $4.3 \times 10^{14} \text{ Hz}$). Different wavelength portions within the visible region evoke different colours in human eye. In Table 4.1, the approximate wavelength range for each of the seven colours is mentioned.

Visible light is very important for us, the invisible forms of electromagnetic radiation are no less important. In fact, these invisible forms of electromagnetic radiation are widely used by us in our daily life particularly for communication and entertainment. For AM radio waves, the electromagnetic waves within frequency range $5.4 \times 10^5 \text{ Hz}$ to $1.6 \times 10^6 \text{ Hz}$ (540 kHz to 1600 kHz) are used, while for FM radio broadcasts frequencies from $8.8 \times 10^7 \text{ Hz}$ to $1.08 \times 10^8 \text{ Hz}$ (88 MHz to 108 MHz) are used. For TV broadcast frequencies from FM band to around 200 MHz are used. Microwaves ($\approx 3 \times 10^9 \text{ Hz}$) are used for cellular phones, radars microwave links, etc. Many cameras have a device that emits a beam of infrared radiation, which on getting reflected from the object to photographed is analyzed to determine the distance and automatically adjusts the focus. Ultraviolet rays (frequencies larger than visible light) can be focused into very narrow beam and are used for many precision applications such as LASIK eye surgery. X-rays are used by dentists and orthopaedics to investigate teeth and bones as the X-rays can pass through flesh but not through teeth and bones. Gamma rays are the shortest wavelength electromagnetic radiation, and they are produced in nature by radioactive materials. Gamma rays are highly energetic and are used to destroy cancer cells by the doctors.

Table 4.1 Range of wavelengths (approx.) of visible light

Colour	λ (in nm)
Violet	400 to 440
Blue	440 to 480
Green	480 to 560
Yellow	560 to 590
Orange	590 to 630
Red	630 to 700

SHORT ANSWER TYPE QUESTIONS

1. What is the direction of force on a charged particle in (a) an electric field and (b) a magnetic field? Explain.
2. What do you mean by 'velocity selector'? Explain its working principle.
3. State Gauss's law.
4. Explain the principle of electrostatic shielding.
5. Obtain an expression for the electric field due to a linear charge distribution.
6. What do you mean by displacement current? Explain.
7. Mention one similarity and one dissimilarity between displacement current and conduction current.
8. Write the four Maxwell's equations and name the law each of them represent.
9. Give the general representation of wave equation and explain it.
10. What do you mean by a plane electromagnetic wave? Explain.
11. Give the important features of electromagnetic waves.
12. Express electric and magnetic field vectors of a sinusoidal electromagnetic plane wave propagating in $+x$ direction in exponential form.
13. Show that the total energy density in electromagnetic wave in free space is $\epsilon_0 E^2$.
14. What is Poynting vector? Define it.
15. What do you mean by radiation pressure? Obtain an expression for it.
16. Write Maxwell's equations for a dielectric medium.
17. Show for a dielectric medium refractive index $n = \sqrt{\epsilon_r}$.
18. Write Maxwell's equations for a conducting medium.
19. What do you mean by skin depth of a conductor? Define it.
20. Give the frequency range of the electromagnetic waves used for FM radio broadcasts.
21. Mention the frequency range of visible light.

LONG ANSWER TYPE QUESTIONS

1. What do you mean by magnetic focusing and magnetic lens? Explain.
2. State and explain Gauss's law. Use this law to obtain an expression for electric field and show its variation with distance.
3. Derive an expression for the electric field due to uniformly charged solid sphere. Hence, show the variation of the field with distance from the centre.
4. Discuss Maxwell's modification of Ampere's law. Hence, show that the modification makes the Ampere's law continuous.
5. Obtain the wave equations for the electric and magnetic field vectors in free space and hence obtain the expression for the wave speed for the electromagnetic waves in free space.
6. What do you mean sinusoidal electromagnetic wave. Obtain the expression for the field vectors of a sinusoidal electromagnetic wave.
7. Show that the intensity of an electromagnetic wave is $\frac{1}{2} \epsilon_0 c E^2$.

8. State and explain Poynting theorem.
9. Obtain wave equations for the electric and magnetic field vectors in a dielectric medium. Hence, show that the speed of electromagnetic waves in a dielectric medium is always less than that in free space.
10. Obtain the wave equation for the electric and magnetic field vectors in a conducting medium. Hence, show that in a conducting medium the field vectors are damped.
11. Write a detailed note on the applications of electromagnetic waves.

NUMERICAL EXERCISES

1. A radio channel in a city broadcasts at a frequency of 830 kHz. At a point some distance from the transmitter the magnetic field amplitude of the electromagnetic wave from the radio channel is 4.82×10^{-11} T. Calculate (a) the wavelength, (b) the wave number, (c) the angular frequency and (d) electric field amplitude. (Ans. (a) 361 m, (b) 0.0174 rad/m , (c) $5.22 \times 10^6 \text{ rad/sec}$, (d) 0.0144 V/m)
2. An electromagnetic wave propagates in a dielectric material. At the frequency of the light, the dielectric constant of the material is 1.74 and the relative permeability is 1.23. If the magnetic field amplitude is 3.80×10^{-9} T, What is the electric field amplitude? (Ans. 0.779 V/M)
3. An electromagnetic wave has an electric field given by $\mathbf{E}(y, t) = -(3.10 \times 10^5 \text{ V/m}) \hat{\mathbf{k}} \sin [ky - (12.65 \times 10^{12} \text{ rad/s}) t]$. In which direction wave is travelling? (b) What is the wavelength of the wave? (c) Write the vector equation for $\mathbf{B}(y, t)$. (Ans. (a) +y = direction, (b) $7.11 \times 10^{-4} \text{ m}$ (c) $\mathbf{B}(y, t) = (-1.03 \times 10^{-3} \text{ T}) \hat{\mathbf{j}} \cos[(8.84 \times 10^3 \text{ rad/m})z - (2.65 \times 10^{12} \text{ rad})t]$.)
4. An electromagnetic wave has a magnetic field given by $\mathbf{B}(x, t) = (8.25 \times 10^{-9} \text{ T}) \hat{\mathbf{j}} \sin [(1.38 \times 10^4 \text{ rad/m})x + \omega t]$. (a) In which direction the wave is propagating? (b) What is the frequency of the wave? (c) Write the vector equation for $\mathbf{E}(x, t)$.
5. An electromagnetic wave with frequency 5.70×10^{14} Hz propagates with a speed of $2.17 \times 10^8 \text{ m/s}$ in a certain piece of glass. Find (a) the wavelength of the wave in the glass, (b) the wavelength of a wave of the same frequency propagating in air, (c) the refractive of the glass for an electromagnetic wave this frequency, (d) dielectric constant for glass at this frequency, assuming the relative permeability is unity. (Ans. (a) 381 nm, (b) 526 nm, (c) 1.38 and (d) 1.91)
6. A monochromatic light source with power output of 60 W radiates light of wavelength 700 nm uniformly in all directions calculate \mathbf{E}_{\max} and \mathbf{B}_{\max} for the 700 nm light at a distance of 5m from the source. (Ans. $12 \text{ V/m}, 4.0 \times 10^{-6} \text{ T}$)
7. An intense light source radiates uniformly in all directions. At a distance of 5 m from the source the radiation pressure on a perfectly absorbing surface is $9.0 \times 10^{-6} \text{ Pa}$. What is the total average power output of the source? (Ans. $8.5 \times 10^5 \text{ W}$)
8. A space probe $2.0 \times 10^{10} \text{ m}$ from a star measures the total intensity of electromagnetic radiation from the star to be $5.0 \times 10^3 \text{ W m}^{-2}$. If the star radiates uniformly in all directions, what is its total average power output? (Ans. $2.5 \times 10^{25} \text{ J}$)
9. A sinusoidal electromagnetic wave from a radio station passes perpendicularly through an open window that has area 0.500 m^2 . At the window, the electric field of the wave has rms value 0.02 V m^{-1} . How much energy does this wave carry through the window during a 30 seconds commercial?

Unit V

5

X-rays Production and Properties

5.1 X-RAYS: PRODUCTION AND PROPERTIES

5.1.1 X-rays: Introduction

X-rays are electromagnetic radiations whose wavelength is in or near 0.001 nm to 1 nm (10^{-12} to 10^{-9} m), which is at least 500 times shorter than visible wavelength. As the shorter wavelengths mean higher photon energies, X-rays photons are much more energetic than visible photons, with energies of 1 eV or more. German physicist Wilhelm Roentgen discovered X-rays in the year 1895. He found that when highly energetic electrons bombard a metallic target, a very penetrating radiation was produced. Unable to identify the radiation, he gave it the name X-rays. Thus, the X-rays emission is the inverse of photoelectric effect.

Since its discovery, X-rays have been used for many applications in medicine and industry. Since the photons of the X-rays are of such high energy, they can penetrate several centimeters inside solids and therefore are used to visualize the interiors of the solids that are opaque to ordinary light, such as broken bones and defects in structural steel. X-rays are used for *computer topography* (or CT scanner). In this process, an X-ray source produces a thin, fan-shaped beam that is detected on the opposite side of the object to be investigated by an array of several hundred detectors in a line. Each detector measures absorption along a thin line across the object. The entire apparatus is rotated around the object in the plane of the beam during a few seconds. The changing photon counting rates of the detectors are recorded digitally. A computer processes the information and reconstruct a picture of absorption over an entire cross-section of the object. In this way tumors and other anomalies can be seen, which is not possible by simple X-ray photographs. X-rays are also being used for bag scan at airports and metro and railway stations.

X-rays have been used for cancer treatment also, as X-rays can damage living tissues. When X-rays photons are absorbed in tissues, their energy breaks molecular bonds and produce highly reactive free radicals, such as H and OH, which in turn can disturb the molecular structure of protein, especially genetic material. Young and rapidly growing cells are particularly susceptible. Thus, X-ray can be used for selective destruction of cancer cells. However, the risk is that a

damaged cell may survive and continues dividing and produce generations of defective cells, causing cancer.

Excessive exposure to X-rays can damage the organism's reproductive system, which may affect its offsprings. Thus, although X-rays are highly useful and over exposure to it is dangerous.

5.1.2 Production of X-rays

Figure 5.1.1 shows the arrangement for the production of X-rays, also called X-ray tube. Two electrodes are enclosed in an evacuated glass tube. Electrons are ejected by thermionic emission from the heated cathode on the left. A potential difference of several thousand volts between the cathode and anode accelerates the electrons toward the anode which acquire kinetic energy of several keV, and hence speeds of $0.1c$ or more. When these electrons strike the metallic anode, the electrons come to an abrupt stop and X-rays are produced. The anode is so placed that X-rays exit in one direction, normally 90° to the electron's path.

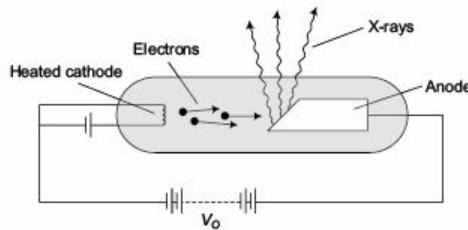


Fig. 5.1.1 Production of X-rays using an X-ray tube.

The ability of the X-rays to penetrate solids of low density was put to medical applications within a few months of their discovery, but it took quite long time to identify what exactly the X-rays are. It was known that an accelerating (or decelerating) charge produces electromagnetic waves. Therefore, it was reasonable to suppose that X-rays were electromagnetic waves produced by the enormous deceleration of the electron stopped by anode. Radiation produced in this manner is called *bremstrahlung*, the German word for breaking radiation. Energy loss due to bremstrahlung is more important for electrons than for heavier particles because electrons are more violently accelerated when passing near nuclei in their paths. The greater the energy of an electron and the greater the atomic number of the nuclei it encounters, the more energetic the bremstrahlung.

5.1.3 X-ray Spectra

Using an X-ray spectrometer it is possible to analyze the distribution of wavelengths produced in an X-ray tube. This kind of distribution is generally recorded as a spectrum, i.e., as *variation of intensity as a function of wavelength or frequency*. Figure 5.1.2 shows two typical X-ray spectra with same accelerating potential but with different anode metals (platinum and molybdenum). The curves exhibits two important features which cannot be explained on the basis of classical theory of bremstrahlung (according to which X-rays should be produced in a broad spread of frequencies, with intensity smoothly varying with frequency). The two important features are:

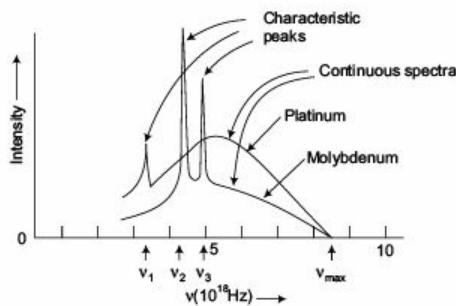


Fig. 5.1.2 X-ray spectra of platinum and molybdenum.

- Both the curves show one or more peaks, on otherwise smooth background, indicating enhanced production of X-rays at certain frequencies. (v_1 for platinum and v_2 , v_3 for molybdenum). Similar peaks appear with the anodes of other metals, at different frequencies that are characteristic of a particular metal. The X-rays at these frequencies are called *characteristic X-rays*. These characteristic X-rays are superposed on continuous X-ray spectrum.
- For a given accelerating potential V , the spectra (with anode of both metals) drop abruptly to zero at a certain maximum frequency ν_{\max} which is the same for both metals. Duane and Hunt found experimentally that ν_{\max} is directly proportional to the accelerating voltage V . According to Duane-Hunt law

$$h\nu_{\max} = K = eV \quad (5.1.1)$$

The second observation can be easily understood on the basis of quantum theory of radiation. The energy of the X-rays are quantized according to $E = h\nu$ and the energy of each quantum is supplied by one of the electrons striking the anode in the X-ray tube. These electrons have kinetic energy $K = eV$. Since the maximum energy an electron can give up is K , no X-ray photon can be produced with energy $h\nu$ greater than K , which explains the existence of a maximum frequency.

X-ray production (except corresponding to the peaks or characteristics X-rays) is an inverse photoelectric effect. The kinetic energy possessed by the electron is transformed into photon energy. Most of the electrons when strike the target undergo numerous glancing collisions, with their energy getting converted into heat. That is why, normally the target anode is made from high melting point metals such as tungsten and sometimes also provided with cooling mechanism. A few electrons, however, lose most or all their energy in single collisions with target atoms. This is the energy that becomes X-rays.

In terms of wavelength, Duane-Hunt law can be expressed as

$$eV = K = h\nu_{\max} = \frac{hc}{\lambda_{\min}} \quad (5.1.2)$$

or

$$\lambda_{\min} = \frac{hc}{eV}$$

Characteristic X-ray Spectra: The continuous X-ray spectrum is the result of the inverse of photoelectric effect, wherein the kinetic energy of the electron is transformed into photon. The characteristic spectra (or line spectra) in form of peaks (or spikes) could be explained only after the advancement of Bohr's theory of atom. A British physicist Henry Moseley explained characteristic X-rays within a few months of the publication of Bohr's theory. Moseley explained, as discussed below, that the characteristic X-rays are produced due to electronic transitions involving inner electrons.

In a multielectron hydrogen like atom, the charge distribution can be approximated to be spherical, and the outer electrons exert no net force on the innermost electrons (as field due to a spherically symmetric shell of charge is zero inside the shell). Therefore, the innermost electron feels only the force of the nuclear charge Ze , and its allowed energy levels E_n is given by

$$E_n = -Z^2 \frac{E_R}{n^2} \quad (5.1.3)$$

where Z is the atomic number and $E_R (= hc R = 13.6 \text{ eV})$ is the Rydberg energy. The term Z^2 implies that in high atomic number atoms, the inner electrons are tightly bound. The atomic transitions, involving inner electrons in the atoms with high Z , involve energies of several thousand eV. Therefore, photon emitted or absorbed in such a transition should be an X-ray photon.

Moseley's explained that in an X-ray tube when the anode is struck by high-energy electrons, one or more electrons are ejected out from the anode. If an electron from an orbit with $n = 1$ is ejected out, an electron from an outer orbit, say $n = 2$, go to the vacant electronic site in $n = 1$ orbit, emitting a photon whose energy E_r is given by (using Eq. 5.1.3)

$$\begin{aligned} E_r &= E_2 - E_1 \\ &= Z^2 R \left(1 - \frac{1}{4}\right) = \frac{3}{4} Z^2 E_R \end{aligned} \quad (5.1.4)$$

Transitions between $n = 2$ and $n = 1$ are identified as K_α transitions. Thus, the K_α photons emitted or absorbed by an atom should have energy $E_r = \frac{3}{4} Z^2 E_R$. Since $E = h\nu$, if frequencies of several K_α X-rays with many anode metals are measured, then (as $\nu \propto Z^2$) a relation

$$\sqrt{\nu} \propto Z \quad (5.1.5)$$

should be obtained. Moseley measured K_α line of more than 20 elements and plotted $\sqrt{\nu}$ against the atomic number Z (Fig. 5.1.3) and found that data fitted into a straight line. This supported Bohr's theory which was very important at that time in addition to establishing the importance of the atomic number.

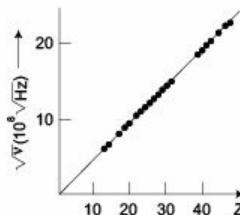


Fig. 5.1.3 $\sqrt{\nu}$ vs. Z

A careful observation of Fig. 5.1.3 shows that the data do not confirm the prediction (5.1.5) exactly. If $\sqrt{v} \propto Z$, the line should pass through the origin, which is not exactly in reality. The line meets the z-axis, close to $Z = 1$. In other words, the data indicates that $\sqrt{v} \propto (Z - 1)$, or equivalently

$$E_r \propto (Z - 1)^2 \quad (5.1.6)$$

This discrepancy was explained by Moseley himself as follows. The prediction that the frequencies of the X-rays is proportional to Z^2 was based on the assumption that the inner electrons feels the force of the nuclear charge Ze and is completely unaffected by the other electrons. This approximation is not completely true because the inner electron does experience some repulsion due to the other electrons. This slightly offsets or screens the attraction of the nucleus. This amounts to a small reduction in the nuclear charge and therefore Ze should be replaced by $(Z - \delta)e$, where δ is a small number close to 1. The energy levels of the inner electron, and hence the X-ray energies should be proportional to $(Z - \delta)^2$ rather than Z^2 . Eq. (5.1.4) therefore should be expressed as

$$E_r = \frac{3}{4} (Z - \delta)^2 E_K \quad (5.1.7)$$

In addition to supporting Bohr's newly formulated atomic model, Moseley's work provided the method to determine the atomic number Z of an element experimentally. This enabled the scientists to establish the correct sequence of the elements in the periodic table. Further, Moseley also found that with an impure anode he could detect the X-rays lines of the impurities. For this reason, he predicted that X-ray spectroscopy would be a useful method of chemical analysis. His prediction has proved true and X-ray spectroscopy is being used extensively. When a sample is put in a beam of electrons, protons or X-rays, inner electrons of the atoms are ejected which then emit X-rays. By measuring the wavelengths (or frequencies), one can identify all elements in the sample present even upto 'trace' level of one part per million or even less.

5.1.4 Crystalline and Amorphous Solids

In solids, the building blocks or units are atoms or molecules. The interacting forces determine properties like electrical conductivity, thermal conductivity, mechanical strength, etc. of a solid. *On the basis of their characteristic properties, the solids are broadly classified into two categories: (1) crystalline solids and (2) amorphous solids.*

1. Crystalline Solids. *The solids having a regular and periodic arrangements of atoms or molecules are called crystalline solids.*

The regular arrangement of its unit (or building block) imparts crystalline solids a definite external geometrical shape.

Few examples of crystalline solids are: quartz, sugar, calcite, metals, mica, diamond, etc.

Some important properties of crystalline solids are given below.

- (i) The constituent units of crystalline solids are arranged in *regular geometrical pattern*.
- (ii) These solids have a *definite external geometrical shape*.
- (iii) Crystalline solids have *long range order* of atoms and molecules, i.e., the specific arrangement of atoms and molecules extend throughout the crystal.

In other words, local arrangements of atoms and molecules are repeated identically at arbitrarily large distances.

We can understand what is meant by *long range order* and *short range order* with the help of Fig. 5.1.4. The figure shows a hypothetical system composed of two types of molecules *A* and *B* or composition A_2B_3 . An atom *A* have three *B* atoms as neighbours, and the atom *B* has two *A* atoms as neighbours, with fixed inter-atomic separations in both cases. There are large number of possible arrangements in which atoms can place themselves in accordance with above conditions. The two possible arrangements in a plane are shown in Fig. 5.1.4. The first [Fig. 5.1.4(a)] is a regular, periodic arrangement of A_2B_3 with the same pattern of atomic positions repeated throughout the plane and we say that the arrangement have a *long range order*, i.e., local arrangement of atoms is repeated identically to arbitrarily large distances. The second arrangement [Fig. 5.1.4(b)] has same local arrangement but no periodicity or regularity, i.e., it does not have long range order but only *short range order*. It is to be noted that the arrangement shown in Fig. 5.1.4 (a) has a short range order besides the long range order.

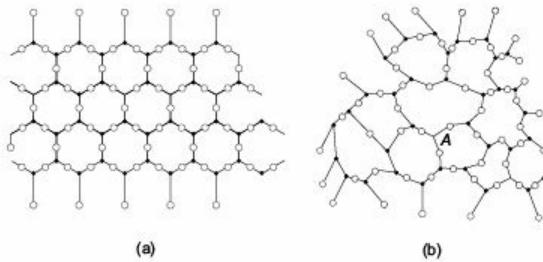


Fig. 5.1.4 Arrangement of two-groups in (a) a crystal and (b) an amorphous material.

- (iv) These solids possess *uniform chemical composition*.
- (v) Crystalline solids are *anisotropic*, i.e., physical properties like electrical conductivity, thermal conductivity, index, etc. are different along different directions (axis). In other words, the bonds of crystalline solids have directional properties.
- (vi) Crystalline solids have a sharp *melting point*, i.e., change from solid to liquid phase occur at a fixed temperature suddenly.
- (vii) Below the temperature of crystallization, the crystalline solids acquire a *stable state*.

2. Amorphous (or Glassy or non-crystalline) Solids. *A solid having no periodic arrangement of atoms (or building blocks or unit cells) is called an amorphous or glassy solid.* Such a solid does not have a fixed atomic/molecular arrangement. Just like liquid, amorphous solids have disordered arrangement of atoms or molecules. However, atoms are not as free to move around as in liquid. Few examples of amorphous solids are: Glass, rubber, plastic, paraffin wax, etc.

Some important characteristic properties of amorphous solids are as follows:

- (i) In amorphous solids constituent atoms or molecules are not arranged in a regular, definite geometrical pattern.
- (ii) Amorphous solids have only short range order. They do not have long range order.
- (iii) Amorphous solids are isotropic, i.e., they have same physical properties in all directions.
- (iv) They do not have a sharp melting point, there is a range of temperature in which they melt.
- (v) They sometimes are considered to be *supercooled liquid*.

5.1.5 X-ray Diffraction: Bragg Law

After the discovery of X-rays by Rontgen in 1895, experiments suggested that they are electromagnetic waves with wavelengths of the order of 10^{-10} m (or 0.1 nm). At about the same time, the idea began to emerge that in a crystalline solid the atoms are arranged in a regular repeating pattern, with spacing between adjacent atoms also of the order of 10^{-10} m. Putting these two ideas together, German physicist Max von Laue proposed in 1912 that a crystal can be used as kind of three-dimensional diffraction grating for X-rays. In other words, a beam of X-rays can be scattered by the individual atoms in a crystal and the scattered waves can interfere just like waves from a diffraction grating. The idea proved correct and Laue and his assistants established that X-rays were waves with wavelengths of order of 10^{-10} m. Since then X-ray diffraction has proved to be an invaluable research tool for measuring X-ray wavelengths and for studying the structure of crystals and complex molecules.

The use of crystals as X-ray diffraction gratings was developed by English physicists, W.L. Bragg and his father W.H. Bragg and is often called *Bragg diffraction* (or scattering or reflection). This technique has been very important since its discovery and is even more important today for the study of crystal structure. To understand the Bragg's method, we consider a crystal as a large number of regularly spaced identical parallel planes, each containing many regularly spaced atoms, as shown in Fig. 5.1.5. Let a plane electromagnetic wave (X-ray) is incident on the atoms which act as scattering centers. The wave induces an oscillating electric dipole moment in each scatter. These dipoles act like little antennas, emitting scattered waves. The scattered waves are in different phase because their distances from the source are different. The scattered wave undergo superposition (or interference). Therefore, the observed resultant intensity (or diffraction maxima) is more in those directions where the waves are in phase.

In Bragg diffraction, the incident direction is traditionally specified by the glancing angle measured up from the plane, rather than the angle down from the normal as in optics. If we consider first waves scattered by atoms in a single plane, all the scattered waves will be in phase in the direction given by the law of reflection (Fig. 5.1.5 (a)), $\theta = \theta'$. This is why Bragg diffraction is also often called Bragg reflection.

Now consider the waves scattered by atoms in two adjacent planes, at a distance d apart (Fig. 5.1.5(b)). As seen in the figure, the path difference for these two waves is $d \sin \theta + d \sin \theta = 2 d \sin \theta$. Thus, the waves from adjacent planes will be in phase if this path difference is an integral multiple of the wavelength, i.e.,

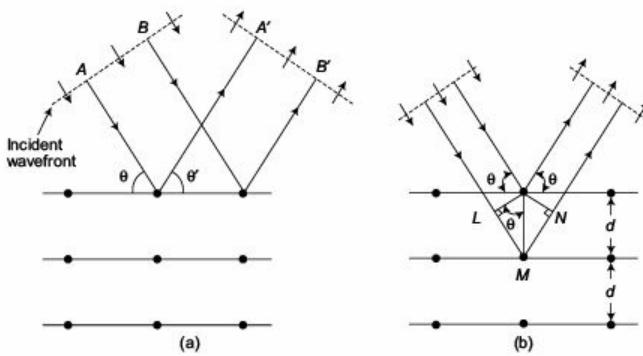


Fig. 5.1.5 Bragg diffraction.

$$2d \sin \theta = n\lambda \quad (5.1.8)$$

where $n = 1, 2, 3, \dots$ and is known as the order of diffraction. In most of the cases, intensity for $n > 1$ is weak and therefore $n = 1$ is used most of the times. Equation (5.1.8) is also known as *Bragg law*. In the directions where Eq. (5.1.8) is satisfied, the waves from all atoms in the crystal will be in phase and a strong maximum will be observed.

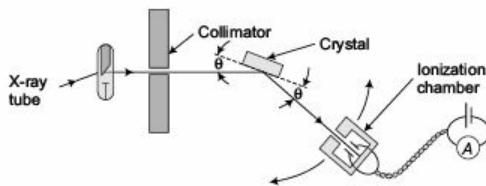


Fig. 5.1.6 An X-ray spectrometer.

Bragg law is used to calculate the spacings d in crystals, using monochromatic X-rays (X-rays of a single wavelength λ). If X-rays of a wide range of wavelengths is used, the wavelengths which satisfy Bragg condition will give maxima. This is used in X-ray spectrometer. A simple X-ray spectrometer, also used by Bragg, is shown in Fig. 5.1.6. The X-rays are passed through a collimator and reflected by a crystal. The intensity of the reflected beam is measured by a detector such as an ionization chamber. By rotating the crystal and detector (in such a manner that angles θ is always equal), the intensity as function of θ is measured which is then used to investigate crystal structures. This use of X-ray is called X-ray crystallography and is an important tool in solid-state physics, materials science and molecular biology.

Example 5.1.1

A potential difference of 10 kV is used to accelerate electrons in an X-ray tube. What is the minimum wavelength of the X-rays produced?

Solution. To produce X-rays with minimum wavelength (or maximum frequency and hence maximum energy), all the kinetic energy of electron must go into producing a single X-ray photon. We know.

$$\begin{aligned}\lambda_{\min} &= \frac{hc}{eV} \\ &= \frac{6.626 \times 10^{-34} \text{ Js} \times 3.0 \times 10^8 \text{ m/s}}{1.602 \times 10^{-19} \text{ C} \times 10.0 \times 10^3 \text{ V}} \\ &= 1.24 \times 10^{-10} \text{ m} = 0.124 \text{ nm}\end{aligned}$$

Example 5.1.2

A beam of X-rays of wavelength 1.54 \AA is incident at certain planes of a silicon crystal. As the angle of incidence is increased from zero, a strong interference maximum is seen at 34.5° . Calculate the interplanar separation. Will you observe other interference maxima at higher incidence angles?

Solution. Bragg equation: $2d \sin \theta = n\lambda$

$$\begin{aligned}\therefore \text{Interplanar separation } d &= \frac{n\lambda}{2 \sin \theta} = \frac{1 \times 1.54 \times 10^{-10} \text{ m}}{2 \times \sin 34.5^\circ} = 1.36 \text{ \AA} \\ \sin \theta &= \frac{n\lambda}{2d} = \frac{n \times 1.54 \times 10^{-10} \text{ m}}{2 \times 1.36 \times 10^{-10} \text{ m}} = n (0.566)\end{aligned}$$

As the value n of 2 or more gives the value of $\sin \theta$ greater than one, which is impossible. Therefore, interference maximum corresponding to higher order will not be seen.

Example 5.1.3

The interplanar spacing in common salt (NaCl) is $d = 0.282 \text{ nm}$. A monochromatic beam of X-rays produces a Bragg maximum when the glancing angle with these planes is $\theta = 7^\circ$. Taking this as the first order maximum ($n = 1$), find the wavelength of the X-rays. What is the minimum required accelerating potential V to produce these X-rays?

Solution. We know, $2d \sin \theta = n\lambda$

$$\begin{aligned}\therefore \lambda &= \frac{2d \sin \theta}{n} = \frac{2 \times 0.282 \times 10^{-9} \text{ m} \times \sin 7^\circ}{1} = 0.069 \text{ nm} \\ \text{From Duane-Hunt law } eV &\geq h\nu = \frac{hc}{\lambda} \\ &= \frac{1240 \text{ eV} \cdot \text{nm}}{0.069 \text{ nm}} = 18000 \text{ eV.}\end{aligned}$$

SHORT ANSWER TYPE QUESTIONS

1. What are X-rays?
2. Explain how X-rays are used for computer topography or CT scan?
3. Mention few applications of X-rays.
4. What is bremsstrahlung? Explain.
5. Name the two types of X-ray spectra and briefly explain them.
6. Explain Duane-Hunt law.
7. Obtain an expression for the existence of λ_{\min} in X-ray spectrum and explain its existence.

8. What do you mean by X-ray spectroscopy? Explain.
9. Differentiate between crystalline and amorphous solids.
10. Why amorphous solids are also called 'glassy' solids?
11. Explain why crystals can act as diffraction grating to X-rays.
12. Express Bragg's law.
13. Discuss the uses of X-ray spectroscopy.

LONG ANSWER TYPE QUESTIONS

1. What are X-rays? Explain giving suitable diagram how are they produced.
2. Write a detailed note on the applications of X-rays and their harmful effects.
3. Explain X-ray spectra detailing the existence of characteristic peaks and a ν_{max} .
4. Discuss Moseley's explanation of characteristics spectrum.
5. Giving characteristics of crystalline and amorphous solids, differentiate between them.
6. Discuss Bragg's X-ray diffraction and its applications.

NUMERICAL EXERCISES

1. The acceleration voltage in X-ray tube is 15 kV. What is the shortest wavelength of the X-rays produced?
(Ans. 0.0837 nm)
2. What is the minimum potential difference between the filament and the target of an X-ray tube if the tube is to produce X-rays of wavelength of 0.150 nm?
3. What is the shortest wavelength produced in an X-ray tube operated with 30 kV?
4. Protons are accelerated from rest by a potential difference of 4kV and strike a metal target. If a photon produces one photon on impact, what is the minimum wavelength of the produced X-rays. Will your answer change if 4 kV electrons are used instead? Why do X-ray tubes use electrons and not protons?
(Ans. 0.310 nm, No)

5.2

Ultrasonics and Acoustics of Buildings

5.2.1 Ultrasonic Waves

The human ear can hear the sound waves in the frequency range 20 Hz to 20 kHz. This range of frequency is known as *audible* range. The sound waves of frequencies above the limits of human audibility, i.e., greater than 20 kHz are called *ultrasonic waves* or *supersonic waves*. However, the term supersonic is nowadays used commonly in aviation for waves having velocities greater than that of sound. Sound waves having frequencies less than 20 Hz are called *infrasonic waves*.

The wavelengths of ultrasonic waves are very small compared to that of audible sound and most of the applications of the ultrasonic waves are due to their small wavelengths. These waves are energetic and show negligible diffraction due to their small wavelength and can be transmitted to longer distance.

5.2.2 Production of Ultrasonic Waves

There are two ways by which ultrasonic waves can be produced. The method actually employed depends upon the required power output and the frequency range.

- (i) **Piezoelectric Ultrasonic Generator:** Before discussing the generator, let us try to understand the piezoelectric effect.

Curie brothers, J. Curie and P. Curie, in 1880, discovered that when mechanical pressure is applied on the opposite faces of certain crystals like quartz, tourmaline, rochelle salt, etc., then electric charges, equal and opposite are developed on the opposite faces of the crystal. (Fig. 5.2.1).

Thus, an electric potential is set up whose magnitude is proportional to the applied pressure. This phenomenon is called *piezoelectric effect*.

The converse piezoelectric effect is also observed, i.e., when a potential difference is applied across the crystal, a change in the crystal dimension is observed. The piezoelectric effect is used to detect the ultrasound while its reciprocal effect is used to generate ultrasound in

which, when alternating potential difference is applied the crystal sets into vibration. A properly cut slice of piezoelectric crystal has a natural frequency of vibration. When the frequency of the applied a.c. signal is equal to the natural frequency of vibration of the crystal, the crystal undergoes resonance and the amplitude of the vibration increases sharply.

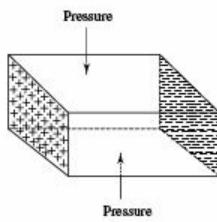


Fig. 5.2.1

Langevin in 1917 devised a method of generating ultrasonic waves by the use of piezoelectric effect, the experimental arrangement for which is shown in Fig. 5.2.2. The high frequency alternating voltage, obtained from Hartley oscillator, is applied to a piezoelectric crystal, Q , placed between metal plates A and B so as to form a parallel plate capacitor with crystal Q as dielectric. The plates are connected to the primary of a transformer which is coupled inductively to the oscillatory circuit of a valve, as shown in Fig. 5.2.2. The frequency of the circuit can be varied by adjusting the variable capacitor C_1 . If the frequency of the oscillatory circuit is equal to the natural frequency of the crystal, resonance occurs and the crystal undergoes linear expansions and contractions at exactly the same rate, i.e., the crystal is set into mechanical vibrations and ultrasonic waves are produced. The ultrasonic waves up to a frequency of 540 kHz can be produced with a moderate size crystal using this method. However, ultrasonic waves of frequencies up to 1.5×10^8 Hz can be produced by using tourmaline crystal.

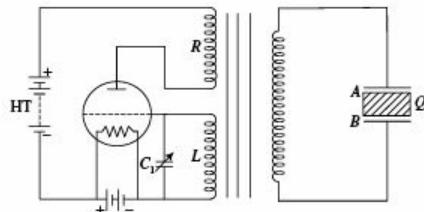


Fig. 5.2.2

(ii) **Magnetostriction Ultrasonic Generator:** Such a generator is based on the magnetostriction effect. Let us try to understand this effect.

Magnetostriction effect: When a bar of ferromagnetic material, such as iron or nickel, is placed in a strong magnetic field parallel to its length, a change (contraction or extension) in

the length of the bar takes place. This effect is called *magnetostriction effect*. The change in length is independent of the sign of the field and depends on the magnitude of the field and the nature of the material.

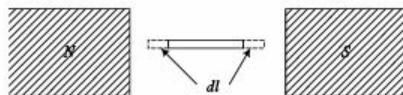


Fig. 5.2.3 Ferromagnetic rod in magnetic field.

If a rod of ferromagnetic material is placed in a rapidly varying magnetic field, the rod alternately expands and contracts with twice the frequency of the applied magnetic field. A varying magnetic field is produced by passing an alternating current through a coil. The rod suffers equal change in length for each half cycle of alternating current, setting up vibrations in the rod whose frequency is twice that of the alternating current. Normally, the amplitude of the vibrations is small. However, if the frequency of the applied a.c. is equal to the natural frequency of the rod, the resonance occurs and the amplitude of the vibrations increase sharply. Sound waves are produced at the ends of the rod. If the applied frequency of the a.c. is in the ultrasonic frequency region, the vibrating rod produces ultrasonic waves. Since the rod vibrates longitudinally. The frequency of the fundamental mode of vibration is:

$$v = \frac{1}{2t} \sqrt{\frac{Y}{\rho}} \quad \dots(5.2.1)$$

where Y is the Young's modulus of the material, ρ is its density and t its thickness or separation between the vibrating ends. Figure 5.2.4 shows the experimental arrangement for producing ultrasonic waves, using magnetostriction oscillator which involves the interaction of a magnetostrictive bar and the oscillations of a valve circuit. The bar is placed inside a coil parallel to its axis.

The bar is normally made up of invar (36% nickel and 64% iron) or monel (65% nickel, 31% copper and 4% Fe, Mn, Si and C). The rod is permanently magnetised before placing it in the coil. When a high frequency a.c. is passed through the coil, the rod is magnetised and demagnetised with the current. The length of the rod varies accordingly and its free ends produce ultrasonic waves. When the frequency of the a.c. becomes equal to the natural frequency of the rod, resonance occurs and the ultrasonic sound waves of maximum amplitude are produced.

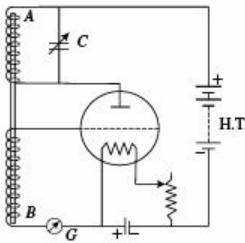


Fig. 5.2.4

5.2.3 Detection of Ultrasonic Waves

We cannot perceive or detect ultrasonic waves directly although some animals, specially the bat can do so. However, we can detect ultrasonic waves by the following methods.

- (i) **Piezoelectric detector:** When a pair of faces of a piezoelectric crystal such as quartz is subjected to ultrasonic waves, opposite electric charges are produced on the other pair of faces, perpendicular to the previous pair of faces. Although the produced charges are relatively small, it can be amplified and detected by some suitable means.
- (ii) **Kundt's tube method:** Kundt's tube can be used to detect ultrasonics in the same way as it is used to detect ordinary sound waves. When ultrasonic waves are passed through the tube, the lycopodium powder sprinkled in the tube collects in the form of heaps at the nodal points and is blown off at the antinodal points.
- (iii) **Thermal detectors:** A probe of fine platinum wire is moved in the region of ultrasonic waves. The temperature of the medium changes due to the formation of alternate regions of compressions and rarefactions. There is a change of temperature at nodes, while at antinodes there is no change in temperature. Due to this, the platinum probe is alternately heated and cooled resulting in the change in resistance. However, no such change occurs at antinode. The change in resistance is detected using a sensitive bridge arrangement.
- (iv) **Sensitive flame method:** When a narrow sensitive flame is moved in a region where ultrasonic waves are present, the flame remains stationary at antinodes and flickers at nodes. This is used to detect ultrasonic waves.

5.2.4 Applications of Ultrasonic Waves

- (i) **Detection of flaws in materials (Non-destructive testing):** To detect flaws or discontinuities in the materials (metals), *pulse echo system* is most commonly employed. In this method, short pulses of ultrasonic waves are transmitted into the material being tested. These pulses get reflected from discontinuities on their path or from any boundary of the material on which they strike. The reflection of pulses also takes place from the back surface of the material. The reflected pulses are picked up by the receiver and are suitably amplified. These pulses are now applied to one set of plates of cathode ray oscilloscope (CRO). The transmitted signal and reflected signal from the flaw and back surface of the material produce a peak each. As shown in Fig. 5.2.5, a transmitting transducer sends a beam of ultrasonics through the test specimen. In the presence of a flaw or a crack or discontinuity in the specimen, the waves are reflected and is detected by the transducer and the corresponding peak on the CRO screen is very weak. The location of the discontinuity with respect to the scanning surface can be obtained by the proper calibration of the CRO time base scale.
- (ii) **SONAR (Sound Navigation and Ranging):** Ultrasonics can be used for the determination of the presence of submerged submarines, rocks, icebergs, etc. or to find the depth of the sea using a system known as Sonar. In this system, a strong beam of ultrasonic waves is transmitted and are picked up on their reflection. The reflection of waves indicates the presence of some reflecting object. The time interval between the transmission and reception of the waves is used to estimate the distance of the reflecting body using the formula $t = 2d/v$, where v is the velocity of the waves and d the distance. The change in frequency of the echo signal due to Doppler effect helps to determine the velocity of the approaching body.

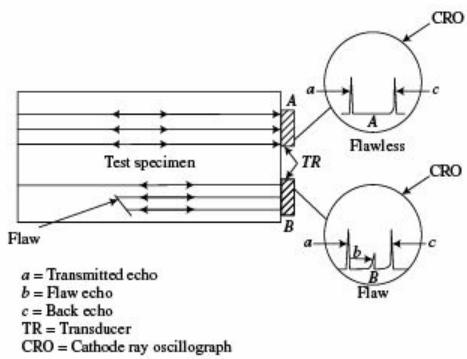


Fig. 5.2.5

- (iii) **Direction signalling:** These waves have been used for directional signalling. On account of their high frequency, these waves can be sent out in the form of a sharp beam and therefore have great advantage over audible sound which spreads out in all directions. It is, therefore, possible to radiate out a large amount of energy in a given direction without the amplitude of the vibrating crystal being too large. Experimentally, it is found that if a vibrating crystal in the form of a circular disc of radius r is used for transmitting the ultrasonic waves, the angle of cone containing the waves is given by $\theta = \sin^{-1} \left(\frac{0.61\lambda}{r} \right)$, which is quite small for short wavelengths.
- (iv) **Ultrasonic mixing:** These waves have been found to be quite useful for forming stable emulsion of immiscible liquids like water and oil by subjecting the mixture to ultrasonic radiations. These waves are also used to liquify gels like aluminium hydroxide in the similar way. Most of the emulsions like polishes, paints, food products and pharmaceutical preparations are prepared by using ultrasonic mixing.
- (v) **Soldering and metal cutting:** These waves are used for soldering, e.g., aluminium cannot be soldered by normal methods. To solder aluminium, ultrasonic waves along with electrical soldering iron is used. These waves are also used for drilling and cutting processes in metals. Ultrasonic welding can be done at room temperature.
- (vi) **Medical applications:** Ultrasonic waves (ultrasound) is used widely for the detection of any abnormal growth, such as tumour, inside the body. It is used to monitor the growth of the unbom baby in pregnant women. The vibration of these waves have a good massage action. The body parts affected by neuralgic or rheumatic pains on being exposed to ultrasonics get relief from pain.
- (vii) **Cleaning and clearing:** These waves are used for cleaning utensils, washing clothes, specially silk fabrics, removing dust and soot from chimney gases. Ultrasonic waves are also used for dispersing fog at airports and for removing water from paper during manufacturing.

- (viii) **Coagulation and crystallisation:** By subjecting the particles of suspended liquid to ultrasonic waves, the particles can be brought quite close to each other so that the coagulation takes place. The crystallisation rate is also affected by ultrasonics. The size of crystals, when molten metal is put to crystallisation, can be made smaller and more uniform by the use of ultrasonics. In metallurgical processes, ultrasonics are used to irradiate molten metals while cooling, so as to refine the grain size, prevent the formation of pores and release the trapped gases.
- (ix) **Destruction of lower life:** Small animals like rats, frogs, fishes, etc. are killed or maimed when exposed sufficiently to the ultrasonic waves. Yeast cells when exposed to these waves lose their ability to reproduce.
- (x) **Thermal effects:** If a fluid is subjected to ultrasonic waves, its temperature rises due to the absorption of waves.
- In addition to the above-mentioned applications, there are numerous other applications of the ultrasonic waves.

Example 5.2.1

A quartz crystal of thickness 1 mm is vibrating in resonance. Calculate the fundamental frequency of the vibration. Given: Young's modulus of quartz $Y = 7.9 \times 10^{10} \text{ Nm}^{-2}$ and its density $\rho = 2.65 \times 10^3 \text{ kg/m}^3$.

Solution.

Here, $t = 1 \text{ mm} = 0.001 \text{ m}$, $Y = 7.9 \times 10^{10} \text{ Nm}^{-2}$, $\rho = 2.65 \times 10^3 \text{ kg/m}^3$. We know, fundamental frequency of vibration is given by:

$$v = \frac{1}{2t} \sqrt{\frac{Y}{\rho}}$$

velocity	$v = \sqrt{\frac{Y}{\rho}}$
wavelength	$\lambda = 2t$

$$\therefore v = \frac{1}{2t} \times \sqrt{\frac{7.9 \times 10^{10}}{2650}} = \frac{5461}{2 \times 0.001} \text{ m} = 2.73 \times 10^6 \text{ Hz}$$

$$= 2730 \text{ kHz.}$$

5.2.5 Acoustics of Building

The branch of science which deals with planning a hall or a building with the objective to provide the best audible sound to the audience is called *acoustic of building* or *Architectural acoustics*.

The credit for this field of science goes to W.C. Sabine, a professor of physics at Harvard University, who in 1911, scientifically studied the problem of satisfactory speech and music in a hall and laid down the following essential features for good acoustics in a hall or auditorium:

- The intensity of sound must be sufficiently large so that it can be distinctly heard in every part of the auditorium and no echoes should be present.
- The quality of the sound must be unaltered, i.e., the relative intensity of the components of a complex sound must be preserved.

- (iii) The successive sounds of speech or music must remain clear and distinct, i.e., there must not be any confusion due to overlapping of sound.
- (iv) The reverberation should neither be too large nor too small. The reverberation time for speech should be between 0.5 to 1 second and that for music between 1 to 2 seconds.
- (v) There should be no concentration of sound in any part of the auditorium.

5.2.6 Reverberation

The sound produced in an auditorium or a hall suffers multiple reflections from the various walls, ceiling and floor of the hall so that in addition to the direct sound, a series of sound waves are heard by the listeners in the hall. The listener feels the persistence of the sound for some time even after the original sound has ceased. This is known as *reverberation*. This is the single most important factor for determining acoustics of a hall. It is defined as *the persistence of audible sound, even after the original sound source has stopped emitting sound waves*. It is due to the following causes.

- (i) The fall in intensity of the sound produced in a hall is exponential and so it takes finite time for the intensity of sound to become inaudible.
- (ii) Due to multiple reflections from the wall, ceiling and floor, the sound reverberates and persists inside the hall for a longer time.
- (iii) As the various sounds reaching a point in the room are not in phase, they undergo interference, causing reinforcement or destruction of sound waves.

Reverberation time: The duration for which sound persists is called reverberation time. It is also defined as *the time taken by the sound wave to fall below minimum audibility; measured from the instant when the source stopped producing sound waves*. Sabine, using an organ pipe of frequency 512 vibrations per second found that its sound became inaudible, when its intensity decreases to one millionth of its intensity just before stopping the organ pipe. Hence, *Sabine defined the standard reverberation time as the time taken by sound to fall to one millionth (10^{-6}) of its intensity just before the source is cut-off*, i.e., the loudness falls by 0 decibels.

Reverberation time depends upon the size of the hall, loudness of the sound and the nature of sound (music or speech) for which the hall is to be used.

5.2.7 Sabine's Experiments

Before 1900, there was no scientific basis for the design of auditoriums for good acoustic properties. The Fogg Art Museum Hall of Harvard University was found to be so defective acoustically that a speaker could hardly make his words audible to the audience. Sabine was entrusted with the responsibility of finding remedial measures and thus began the study of planned acoustics for auditoriums.

Sabine found that in the hall, when empty, a sound persisted for about 5.5 seconds after the source of sound stopped producing sound waves. Although, this time was less when the hall was full, it was too large to be comfortably audible. As mentioned earlier, the time of reverberation T was taken as the time in which the sound intensity fell to 10^{-6} of its original value. Sabine was able to reduce the time of reverberation to 1 second by putting cushions which absorbed the sound. He found that T is inversely proportional to total absorption, i.e.,

$$T \propto \frac{1}{A}$$

He also found that T is directly proportional to the volume of the hall, i.e.,

$$T \propto V.$$

Combining the above two expressions, we get:

$$T \propto \frac{V}{A}$$

or,
$$T = \frac{kV}{A}$$

where k is a constant.

If a_i is the absorption coefficient per unit area and S the surface area of the absorbing material, then

$$T = \frac{kV}{\Sigma aS} = \frac{kV}{A} \quad \dots(5.2.2)$$

The value of k was determined by Sabine to be 0.163, so the Sabine's formula for standard time of reverberation is given by:

$$T = \frac{0.163 V}{A} = \frac{0.163 V}{\Sigma aS} \quad \dots(5.2.3)$$

5.2.8 Derivation of Sabine's Formula for Reverberation Time

It is assumed in the derivation of this formula that there is a uniform distribution of sound energy in the hall. This is a reasonable assumption because when a source starts producing sound in a hall, the reflections at the walls produce a sound energy distribution that becomes progressively uniform with increasing time. Except in the vicinity of the source, the sound energy distribution may be assumed to be completely uniform and to have a direction of flow which is essentially random.

First we calculate the *energy density* E , i.e., *the rate at which energy falls on the wall and other surfaces* and hence the rate at which it is being absorbed. Then, we shall obtain an expression for the final steady value of E in terms of the rate of emission or power P of the source. Finally, we shall formulate and solve a differential equation on the principle that the rate of change of E at any instant is determined by the excess P over the total absorption.

Consider the reception of sound energy by a small element of area ds of a plane wall AB (Fig. 5.2.6) From the centre of ds , two circles of radii r and $r + dr$ are drawn such that they lie in the plane containing the normal to the element ds . Consider the portion of the area between the circles and also lying between the angles θ and $\theta + d\theta$ which the two radii make with the direction of the normal at the surface ds , the shaded portion in the figure. The arc of this area is $rd\theta$ and the radial length is dr .

$$\therefore \text{Area} = rd\theta dr$$

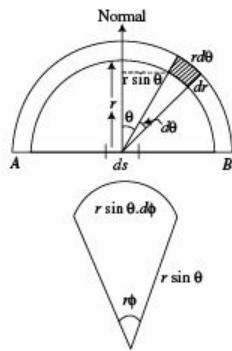


Fig. 5.2.6

If the whole figure is rotated about the normal through an angle $d\phi$ (radius of the rotating figure being $r \sin \theta$), the distance travelled by this area will be the circumferential distance $r \sin \theta d\phi$.

\therefore Volume traced by this element area

$$= r d\theta dr (r \sin \theta d\phi) = r^2 \sin \theta d\theta dr d\phi$$

Multiplying this volume element by energy density E , we get

$$\text{Energy contained in this volume} = Er^2 \sin \theta d\theta dr d\phi.$$

This sound energy is travelling through the element equally in all directions. Therefore, the energy travelling per unit solid angle in any direction

$$= \frac{Er^2 \sin \theta d\theta dr d\phi}{4\pi}$$

The solid angle subtended by ds at the element of considered volume

$$= \frac{ds \cos \theta}{r^2}$$

\therefore Energy travelling towards ds from the considered volume element

$$\begin{aligned} &= \frac{Er^2 \sin \theta d\theta dr d\phi}{4\pi} \cdot \frac{ds \cos \theta}{r^2} \\ &= \frac{E.ds}{4\pi} dr \sin \theta \cos \theta d\theta d\phi \end{aligned} \quad \dots(5.2.4)$$

To find the energy received by ds in unit time, the above expression is integrated from $r = 0$ to $r = C$, where C is velocity of sound. Integrating Eq. (5.2.4), first with respect to ϕ , we get

$$\therefore \text{Energy received by } ds = \frac{E.ds}{4\pi} dr \sin \theta \cos \theta d\theta \int_0^{2\pi} d\phi$$

$$= \frac{E.ds}{2} dr \sin \theta \cos \theta d\theta \quad \left(\because \int_0^{2\pi} d\theta = 2\pi \right)$$

Now integrating w.r.t. θ , we have

$$\begin{aligned} \text{Energy received by } ds &= \frac{Eds}{2} dr \int_0^{\pi/2} \sin \theta \cos \theta d\theta \\ &= \frac{Eds}{4} dr \int_0^{\pi/2} \sin 2\theta d\theta \quad (\because 2 \sin \theta \cos \theta \sin 2\theta) \\ &= \frac{Eds}{4} dr \quad \left(\because \int_0^{\pi/2} \sin 2\theta d\theta = 1 \right) \end{aligned}$$

Now integrating w.r.t. r from 0 to C , we get

Energy received at ds per unit time

$$= \frac{Eds}{4} \int_0^C dr = \frac{EdsC}{4}$$

If a is the absorption coefficient of the wall AB whose element ds is a part, then

$$\text{Energy absorption by } ds \text{ per unit time} = \frac{EC ds a}{4}$$

Total absorption by the surfaces of the wall where sound is falling

$$= \frac{EC}{4} \sum a ds = \frac{ECA}{4} \quad \dots(5.2.5)$$

where $\sum a ds = A$ is the total absorption on all the surfaces on which sound is falling.

Let P be the power output, i.e., rate of emission of sound energy from the source and V the total volume of the hall. Then, when the energy density is E , the total sound energy in the hall is EV .

\therefore Rate of growth of sound energy

$$= \frac{d}{dt}(EV) = V \frac{dE}{dt}$$

Also, the rate of growth of energy

$$\begin{aligned} &= \text{Rate of energy generation by the source} \\ &\quad - \text{rate of absorption of energy by surfaces} \end{aligned}$$

$$\text{or, } V \frac{dE}{dt} = P - \frac{ECA}{4} \quad \dots(5.2.6)$$

When the steady state is attained, $\frac{dE}{dt} = 0$, and the value of E at this instant is denoted by E_m , then

$$0 = P - \frac{E_m CA}{4}$$

or, $E_m = \frac{4P}{CA}$... (5.2.7)

From Eq. (5.2.6), we have

$$\frac{dE}{dt} = \frac{P}{V} - \frac{CA}{4V} \cdot E$$

or, $\frac{dE}{dt} = \frac{4P}{CA} \alpha - \alpha E$ $\left(\text{Putting } \frac{CA}{4V} = \alpha \text{ so that } \frac{1}{V} = \frac{4\alpha}{CA} \right)$

or, $\frac{dE}{dt} + \alpha E = \frac{4P}{CA} \alpha$

Multiplying both sides by $e^{\alpha t}$, we get

$$\left(\frac{dE}{dt} + \alpha E \right) e^{\alpha t} = \frac{4P}{CA} \alpha e^{\alpha t}$$

or, $\frac{d}{dt}(Ee^{\alpha t}) = \frac{4P\alpha e^{\alpha t}}{CA}$

Integrating, we get

$$Ee^{\alpha t} = \frac{4P}{CA} e^{\alpha t} + K \quad \dots (5.2.8)$$

where K is a constant of integration, whose value is determined using boundary conditions.

Growth of energy density: If t is measured from the instant when the source starts emitting the sound energy, then at $t = 0$, $E = 0$. Putting this in Eq. (5.2.8), we get

$$K = -\frac{4P}{CA}$$

Substituting this value in Eq. (5.2.8) we get

$$\begin{aligned} Ee^{\alpha t} &= \frac{4P}{CA} e^{\alpha t} - \frac{4P}{CA} \\ E &= \frac{4P}{CA} (1 - e^{-\alpha t}) \\ E &= E_m (1 - e^{-\alpha t}) \end{aligned} \quad \dots (5.2.9)$$

The expression shows the growth of E with time t . The growth is exponential, as shown in Fig. 5.2.7, at $t = \infty$, $E = E_m$.

Decay of energy density: If the source is cut-off when E has reached its maximum value E_m , so that $P = 0$ and $t = 0$, then

$$E = E_m = \frac{4P}{CA}$$

From Eq. (5.2.7), $K = E_m$, therefore, we have

$$\begin{aligned} Ee^{\alpha t} &= E_m \\ E &= E_m e^{-\alpha t} \end{aligned} \quad \dots(5.2.10)$$

This expression shows the decay of energy density with time after the source is cut-off. The decay is exponential and is shown in Fig. 5.2.8.

The Sabine's formula or the expression for standard reverberation time T is obtained as follows. The reverberation time is the time in which energy density decays to one-millionth (10^{-6}) of its value just before the source is cut-off, i.e., when $t = T$,

$$\begin{aligned} E &= \frac{E_m}{10^6} \\ \text{or, } \frac{E}{E_m} &= 10^{-6} \end{aligned}$$

From Eq. (5.2.10), we have

$$e^{-\alpha T} = 10^{-6} \Rightarrow e^{\alpha T} = 10^6$$

Taking log on both sides,

$$\alpha T = 6 \log_e 10 = 2.3026 \times 6$$

$$\text{or, } \frac{CA}{4V} T = 2.3026 \times 6$$

$$\text{or, } T = \frac{4 \times 2.3026 \times 6 \times V}{340 \times A} \quad (\text{Taking } C = 340 \text{ m/s})$$

$$\text{or, } T = \frac{0.163V}{A} = \frac{0.163V}{\Sigma aS} \quad \dots(5.2.11)$$

This is the Sabine's formula for standard reverberation time and is in good agreement with experimental results.

Jaeger's method for derivation of Sabine's formula: The absorption coefficient α for the surface of a material is defined as the ratio of the sound energy absorbed by the surface to the sound energy absorbed by a perfect absorber, such as an open window. The unit of absorption coefficient is Sabine. So the effective absorbing area A for a surface having total area S and absorption coefficient a is equal to aS , i.e., $A = aS$.

If a_1, a_2, a_3, \dots are the absorption coefficients for each type of reflecting surface in the hall and S_1, S_2, S_3, \dots the corresponding surface areas, then the average value of absorption coefficient a is given by

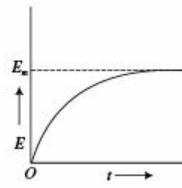


Fig. 5.2.7 Growth of energy density.

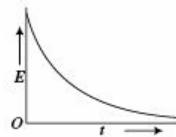


Fig. 5.2.8 Decay of energy density.

$$a = \frac{a_1 S_1 + a_2 S_2 + a_3 S_3 + \dots}{S_1 + S_2 + S_3 + \dots} = \frac{\sum a_i S_i}{S}$$

If we consider I to be the average intensity, i.e., intensity per unit volume of the hall, at any instant of time and dI the fall in intensity due to absorption in a small interval of time dt , then

$$dI = -a \cdot n \cdot I \cdot dt \quad \dots(5.2.12)$$

where, n corresponds to the number of reflections of sound waves per second and a is the absorption coefficient of all absorbing surfaces.

By statistical method, it was shown by Jaeger that sound travels an average distance $\frac{4V}{S}$ between two successive reflections.

Time taken between two successive reflection = $\frac{4V}{SC}$, where C is the velocity of the sound wave.

$$n = \frac{SC}{4V} \quad \dots(5.2.13)$$

Substituting this value in Eq. (5.2.12), we get

$$dI = -a \cdot \frac{SC}{4V} \cdot I \cdot dt$$

$$\text{or, } \frac{dI}{dt} = -a \cdot \frac{SC}{4V} \cdot I$$

$$\text{or, } \frac{dI}{I} = -a \cdot \frac{SC}{4V} \cdot dt \quad \dots(5.2.14)$$

If I_0 is the steady intensity at the instant when the source of sound is cut-off and I_t its value after t seconds, then integrating the above expression, we have

$$\int_{t_0}^t \frac{dI}{I} = - \int_0^t \frac{aSC}{4V} \cdot dt$$

$$\text{or, } \log_e \frac{I_t}{I_0} = -\frac{aSC}{4V} \cdot t \Rightarrow \frac{I_t}{I_0} = e^{-\frac{aSC}{4V} t}$$

$$\text{By the definition for reverberation time } T, \frac{I_t}{I_0} = 10^{-6}$$

$$\therefore e^{-\frac{aSC}{4V} T} = 10^{-6}$$

$$\text{or, } \frac{aSC}{4V} T = 6 \log_e 10 = 6 \times 2.3026 \times 1 = 13.8156$$

Taking velocity of sound

$$C = 340 \text{ m/s}$$

$$\therefore \text{Reverberation time}, \quad T = \frac{13.8156 \times 4V}{aS \times 340} = \frac{0.163V}{aS}$$

This is Sabine's formula for reverberation time.

5.2.9 Importance of Reverberation Time–Optimum Reverberation Time

For good acoustics in a hall, the reverberation time should have some optimum value, i.e., it should neither be too large nor too small. If the reverberation time is too small, the sound vanishes almost instantaneously and gives the hall *dead effect*. If it is too large, sound will undergo multiple reflections and overlappings, thereby making the sound inaudible and causing confusion to the listeners. Therefore, knowing the absorption coefficients of the various surfaces in a proposed hall, the optimum reverberation time of the proposed hall should be calculated in advance. For an existing hall, we can have any desired reverberation time by suitably choosing the surface materials. It is therefore important to know the optimum value of reverberation time which depends upon the purpose for which hall is constructed. For the halls where music is to be played slightly longer reverberation time is required than that for the halls where a speech is to be made.

Sabine's method of calculating reverberation time is considered to be the most standard one. For the determination of reverberation time, he took an organ pipe blown at a definite frequency (512 Hz) and constant pressure. The instant of cutting off the sound and the instant when it became inaudible was recorded on a chronograph drum. An accuracy of $\pm 0.05\text{s}$ was claimed by this method. After performing a large number of measurements, Sabine made the following conclusions:

- (i) Reverberation time is independent of the position of the source and the listener in the hall.
- (ii) The effect of any given amount of absorbent is independent of its position in the hall.

The optimum reverberation time (in seconds at 500 Hz) for various hall volumes is given in Table 5.2.1.

Table 5.2.1 Optimum reverberation time

Volume of Hall (m^3)	2.83×10^2	4.25×10^2	5.66×10^2	8.9×10^2
Choir	1.0	1.0	1.2	1.25
Orchestra	0.8	0.85	0.9	1.00
Speech or music	0.6	0.65	0.7	0.75

5.2.10 Measurement of Absorption Coefficient

The coefficient of absorption of a material is defined as the ratio of the sound energy absorbed by the surface to that of the total incident sound energy on the surface.

Absorption coefficient

$$a = \frac{\text{Sound energy absorbed by the surface}}{\text{Total sound energy incident on the surface}}$$

An open window is considered to be a perfect absorber as all sound waves falling on an open window pass through. Hence, the standard unit of absorption is taken as the unit area of an open window. Thus, the absorption coefficient of a material is defined as the ratio of sound energy absorbed by a certain area of the surface to that of a window of the same area. In other words, the

absorption coefficient of a surface is defined as the reciprocal of its area which absorbs the same sound energy as absorbed by a unit area of an open window. For example, if 10 m^2 of a certain material absorbs the same sound energy as absorbed by 1 m^2 of an open window, then

$$\text{Absorption coefficient } a = \frac{1}{10} = 0.1$$

The unit of absorption coefficient is Sabine or O.W.U. (Open window unit). The absorption coefficients of some materials are given in Table 5.2.2.

Table 5.2.2 Absorption coefficient of some materials

Material	<i>a</i>	Material	<i>a</i>
Glass	0.02	Carpets	0.15 – 0.30
Lime plaster	0.02 – 0.04	Chairs	0.17 – 0.21
Marble	0.01	Acoustic plaster	0.27
Brick wall (30 cm thick)	0.03	Audience per person	0.4367
		Cushions (75 cm thick)	1.5 – 1.8

Measurement of absorption coefficient: If the absorption coefficient of the material of a surface is *a*, the fraction $(1 - a)$ of the sound energy incident upon the surface is reflected back from the surface. The following two methods are used for the determination of the absorption coefficient of a material:

- (i) The first method is based on the determination of standard times of reverberation in the hall without and with a standard large sample of the material whose absorption coefficient is to be determined.

First, the reverberation time T_1 of a sound wave of certain frequency is determined in the hall without the material whose absorption coefficient is to be measured. If $A = \sum a_i S_i$ is the absorption due to the walls of the hall, then using Sabine's formula, we have:

$$T_1 = \frac{0.16V}{A} = \frac{0.16V}{\sum a_i S_i} \quad \dots(5.2.15)$$

Then the material whose absorption coefficient is to be determined is placed in the hall. Suppose *S* is surface area of the material and *a'* its absorption coefficient. The reverberation time T_2 of the same wave is determined again.

$$\text{So that, } T_2 = \frac{0.16V}{A + a'S} \quad \dots(5.2.16)$$

Taking reciprocal and then subtracting Eq. (5.2.15) from Eq. (5.2.16), and rearranging, we get

$$\frac{1}{T_2} - \frac{1}{T_1} = \frac{a'S}{0.16V}$$

$$\text{or, } a' = \frac{0.16V}{S} \left[\frac{1}{T_2} - \frac{1}{T_1} \right] \quad \dots(5.2.17)$$

Thus, knowing V, S, T_1 and T_2 , we can determine a' .

- (ii) In this method, the reverberation times of two sources emitting powers P_1 and P_2 are determined. Even if the absolute values of P_1 and P_2 are not known, it is sufficient to know their ratio. The steady state energy densities of these sources, as we know, are $\frac{4P_1}{CA}$ and $\frac{4P_2}{CA}$. During the decay, they would reach the minimum audible value E_0 in times T_1 and T_2 (by definition), given by

$$E_0 = \frac{4P_1}{CA} e^{-\alpha T_1}$$

and

$$E_0 = \frac{4P_2}{CA} e^{-\alpha T_2}, \quad \text{where } \alpha = \frac{CA}{4V}.$$

Dividing the above two expressions, we get

$$\frac{P_1}{P_2} = e^{\alpha(T_1 - T_2)}$$

Taking logarithm of both sides, we have

$$\alpha = \frac{\log_e(P_1/P_2)}{(T_1 - T_2)} = \frac{CA}{4V}$$

$$\text{or, } A = \frac{4V \log_e\left(\frac{P_1}{P_2}\right)}{C(T_1 - T_2)}$$

As,

$$\therefore A = aS \quad \dots(5.2.18)$$

Thus, knowing $V, P_1/P_2, S, C, T_1$ and T_2 , we can calculate a .

5.2.11 Factors Influencing Building Acoustics

In an acoustically good hall, every syllable or musical note should reach every point of the hall at an audible level and then should quickly fade away to make room for the next syllable or group of notes. If it is not so, acoustically the hall is defective. The following factors influence the acoustic of a hall.

- (i) **Reverberation time:** This is the most important factor for the acoustics of a hall. We have discussed this factor earlier. A hall should have the optimum reverberation time, as determined by Sabine formula. It depends upon the size of the hall, loudness and nature of the sound for which the hall is used. For a frequency of 512 Hz, a reverberation time in the range 1 to 1.5 sec

is considered to be good for a small hall while for a large one, it lies between 2 and 3 sec. The reverberation can be controlled by the following factors.

- (a) Windows and ventilators in the hall can be used to control reverberation time by opening or closing them.
- (b) Curtains with proper foldings.
- (c) Covering the floor with carpets.
- (d) Number of persons (audience) present in the hall.
- (e) Providing acoustic tiles.
- (f) Decorating walls with suitable absorbent materials such as wall pictures, glass wool, fibre-board, etc.
- (ii) **Loudness of sound:** Normally to control reverberation, certain sound absorbers are used which weaken the intensity of sound. Therefore, sufficient loudness of sound is essential so that it can be heard in all parts of the hall. Placing a parabolic reflector at the back of the speaker enables a speaker to be heard at a large distance if the speaker is at the focus of the reflector. Also low ceiling reflects sound towards the audience.
- (iii) **Echelon effect:** Any regular spacings like steps or railings, if present in a hall, reflect sound which produce echoes with regular phase difference. This produce a musical note which is heard along with the original sound by the listeners, causing confusion. This is called *echelon effect*. This can be avoided by having irregular spacing or steps of unequal width or covering them with sound absorbing material.
- (iv) **Focussing due to walls and ceilings:** If there is focussing of sound waves due to focussing surfaces (such as concave, spherical, cylindrical, parabolic walls or ceilings) in the hall, sound is concentrated in a particular region while in some other region, there may not be enough sound to be heard clearly. To avoid this, there should be no curved surfaces and if present, they should be covered with suitable sound absorbing materials.
- (v) **Echoes:** An echo is heard when direct and reflected sound waves reach a listener with a time interval of about 1/7 second. The reflected sound, if arrives earlier than this, raises the loudness of the original sound, while those arriving later produce echoes. Echoes can be avoided by covering the long distance walls and the high ceiling with suitable sound absorbing materials.
- (vi) **Extraneous noise and sound insulation:** Noise or unwanted sound should not enter an auditorium. Also, the sound should not leave the auditorium. The prevention of sound from leaving or entering the hall is called *sound insulation* or *sound proofing*.

Example 5.2.2.

An auditorium has a volume of 9000 m^3 . Reverberation time of 1.4 second is required in it. What should be the total absorption in the hall?

Solution. Here,

$$V = 9000 \text{ m}^3, T = 1.4 \text{ s}$$

We know,

$$T = \frac{0.163V}{aS}$$

$$\text{Total absorption } aS = \frac{0.163V}{T} = \frac{0.163 \times 9000}{1.4} = 1048 \text{ O.W.U.}$$

Example 5.2.3

A lecture hall with a volume of 1500 m^3 is formed to have a reverberation time of 1.3 sec. What is the total absorbing power of all the surfaces in the hall? If the area of the sound absorbing surface is 280 m^2 , what is the average absorption coefficient?

Solution. Here,

$$V = 1500 \text{ m}^3, \quad T = 1.35 \text{ sec}, \quad S = 280 \text{ m}^2$$

We know,

$$T = \frac{0.163V}{\Sigma aS} = \frac{0.163V}{A}$$

$$\text{Total absorbing power: } \Sigma aS = \frac{0.163V}{T} = \frac{0.163 \times 1500}{1.35} = 181.11$$

Average absorption coefficient:

$$\bar{a} = \frac{181.11}{280} = 0.65.$$

Example 5.2.4

The volume of a hall is 1250 m^3 . The wall area is 230 m^2 , the floor area is 125 m^2 and the ceiling area is 125 m^2 . The average sound absorption (i) for walls is 0.04, (ii) for the ceiling is 0.85, and (iii) for the floor is 0.06. Find the average absorption coefficient and reverberation time.

Solution.

The average absorption coefficient

$$\begin{aligned} \bar{a} &= \frac{a_1S_1 + a_2S_2 + a_3S_3}{S_1 + S_2 + S_3} \\ &= \frac{0.04 \times 230 + 0.85 \times 125 + 0.06 \times 125}{230 + 125 + 125} = 0.256 \end{aligned}$$

Total absorption in the hall

$$\begin{aligned} &= \bar{a}S = 0.256 \times (230 + 125 + 125) = 0.256 \times 480 \\ &= 123 \text{ Sabines} \end{aligned}$$

$$\text{Reverberation time } T = \frac{0.163V}{\bar{a}S} = \frac{0.163 \times 1250}{123} = 1.66 \text{ seconds.}$$

SHORT ANSWER TYPE QUESTIONS

1. Give the frequency range for infrasonic, audible and ultrasonic waves.
2. What is piezoelectric effect? How can this effect be used to produce and detect ultrasonic waves?

3. Discuss magnetostriction effect and explain how it can be used to produce ultrasonic effect?
4. Briefly explain the ways to detect ultrasonic waves.
5. How can ultrasound waves be used for non-destructive testing?
6. What is SONAR? Explain its principle of working.
7. What is ultrasound? Discuss its medical applications.
8. Mention the essential features of good acoustics of a building.
9. What is reverberation? Define reverberation time.
10. Define absorption coefficient and give its unit.
11. Explain the importance of reverberation time for the good acoustics of a building. Hence, define 'optimum reverberation time' and 'dead effect'.
12. What do you mean by focussing of sound waves? How this effect can be used for better acoustic?
13. Mention different ways of controlling reverberation time of a hall.
14. What is echelon effect? How can it be avoided?

LONG ANSWER TYPE QUESTIONS

1. Discuss various ways of producing ultrasonic waves. How can ultrasonic waves be detected?
2. Explain various applications of ultrasonic waves.
3. Discuss Sabine experiment and obtain an expression for reverberation time.
4. Derive Sabine's formula for reverberation time of a hall and show the growth and decay of energy density.
5. Obtain an expression for reverberation time using Jaeger's method.
6. Define absorption coefficient. Discuss the two methods for its determination.
7. What are various factors that influence the acoustics of hall? What are the remedial measures?

NUMERICAL EXERCISES

1. An auditorium has a volume of 120000 m^3 . It has reverberation time of 1.5 sec. What is the average absorbing power of the surfaces, if the total sound absorbing surface is 25000 m^2 ?
[Ans. 1.512 unit]
2. The volume of a room is 600 m^3 . The wall area of the room is 220 m^2 , the floor area is 120 m^2 and the ceiling area is 120 m^2 . The absorption coefficients for the walls, the ceiling and the floor area are 0.03, 0.80 and 0.06 respectively. Calculate the reverberation time.
[Ans. 0.885s]
3. The time of reverberation of an auditorium with and without 500 audience is 1.5s and 1.4s respectively. Calculate the reverberation time with 1000 persons in the auditorium.
[Ans. 1.31s]
4. A lecture hall of volume $1.2 \times 10^5 \text{ m}^3$ has a total absorption of 13200 m^2 of O.W.U. Entry of people into the hall raises the absorption by another 13200 m^2 of O.W.U. Find the change in reverberation time.
[Ans. 0.736s]
5. The time of reverberation of an empty hall is 1.5 sec. With 500 persons in the hall, the time of reverberation falls to 1.4 sec. Find the number of persons present in the hall, if the time of reverberation falls down to 1.312 sec.
[Ans. 1000 persons]

Unit VI

Physics of Materials

6

Physics of Semiconductors

6.1 FORMATION OF ENERGY BANDS IN SOLIDS

The energy of electrons in an atom cannot have any arbitrary value but only some definite values in accordance with quantum mechanical laws. In a free atom, the energy levels are well-defined. However, for an atom in a solid, where it is surrounded by other atoms, the energy levels get modified. This modification is not significant in the energy levels of the electrons in the inner shells. But the energy levels of the electrons in the outer shells, particularly in the outermost shell, get appreciably modified. This is because the electrons in the outermost shells are shared by more than one atom in the solid.

In order to understand the modification of energy levels, we consider silicon crystal having N atoms. The electronic configuration of a silicon atom is $1s^2 2s^2 2p^6 3s^2 3p^2$. The outer shell contains 2 electrons in $3s$ subshell and 2 electrons in $3p$ subshell. In the outermost shell $2N - 3s^2$ energy levels are completely filled while $6N - 3p^6$ energy levels are only partially occupied by $2N$ electrons, i.e., only $2N$ energy levels are filled.

Figure 6.1 shows the energy levels of silicon atom in the crystal. The actual interatomic separation is d_0 , i.e., $r = d_0$. To understand that how energy levels get modified, let us imagine the atoms are brought near to each other to form the crystal such that finally the interatomic separation reduces to d_0 . We discuss energy levels at four interatomic separations.

- (i) At $r = d_3$. At this interatomic separation, electrons in one atom do not get influenced by the other atoms. Therefore, each N atom retains its energy levels unperturbed.
- (ii) At $r = d_2$. On reducing the interatomic separation to d_2 , the interaction between the electrons in the outermost shell becomes appreciable. As a result, $3s$ and $3p$ energy levels of the atoms get slightly changed and instead of isolated energy levels, closely packed energy levels result. Now, the $2N$ s -levels will not have same energy but will be spread in small energy band. Similarly, $6N$ p -levels will also spread into a small energy band. This spreading of energy reduces the energy gap between $3p$ and $3s$ levels compared to that in an isolated atom. Because there are approximately 10^{29} atom/m³ the energy levels due to spreading of $3s$ and $3p$ levels are very closely spaced. *The collection of closely spaced energy levels is referred to as energy band.* In each band, energy levels are very closely spaced, the spacing

between successive levels is of order of $\left(\frac{\hbar^2}{mV^{2/3}}\right)$, where $\hbar = \frac{h}{2\pi}$ is Planck's constant, m the electronic mass and V the volume of the solid.

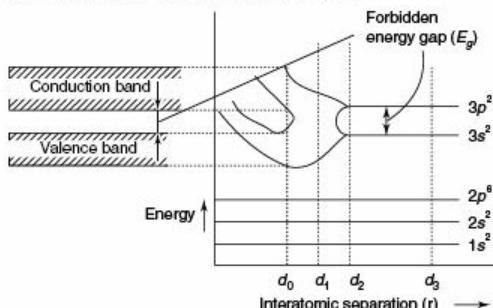


Fig. 6.1

- (iii) At $r = d_1$. On further reduction of interatomic separation to d_1 , the energy gap between $3s$ and $3p$ levels completely disappears and the $8N$ levels ($2N$ s -level and $6N$ p -level) are continuously distributed forming an energy band. In such a situation, it is not possible to distinguish between the electrons belonging to $3s$ and $3p$ subshells. It can only be said that $4N$ levels are filled and $4N$ levels are empty.
- (iv) At $r = d_0$. When the interatomic separation reduces to the equilibrium distance in the actual crystal ($r = d_0 \equiv 1 \text{ \AA}$), the band of $4N$ filled energy levels and the band of $4N$ empty levels get separated by an energy gap called the forbidden energy gap, E_g . The lower completely filled band is called the *valence band* and the upper empty energy band is called the *conduction band*.

Depending on the magnitude of the forbidden energy gap E_g , solids are classified as metals, semiconductors and insulators. We discuss below this classification in detail.

6.2 DISTINCTION BETWEEN METALS, INSULATORS AND SEMICONDUCTORS

Metals are good conductors of electricity, insulators are very poor conductors of electricity, and the conductivity of semiconductors lies in between that of metals and insulators.

Fermi level is a hypothetical level in the energy levels of a crystal (corresponding energy is Fermi energy) all energy levels below which are completely occupied at zero Kelvin and levels above it completely empty. As the temperature is raised above zero Kelvin, some of the electrons acquire enough thermal energy to cross over the Fermi level and jump to conduction band. The electrons in conduction band are free to move inside the crystal and the movement of these free electrons constitute the electric current. The magnitude of the forbidden energy gap determines the number of free electrons available in the conduction band and hence conductivity of a solid at a particular temperature. From the electrical conductivity point of view solids are generally classified as metals, semiconductors and insulators, discussed below (we do not include superconductors as they are special class of materials and also at present superconductors are not available at room temperature).

- (i) **Metals:** The energy band structure of a metal is shown in Fig. 6.2. There are two possible arrangements of energy bands which leads to metallic behaviour of a crystal. In the first kind, an empty conduction band overlaps the completely filled valence band [Fig. 6.2(a)]. Such a metal is called a *band overlap metal*. The second possibility is that *conduction band is partially filled* [Fig. 6.2 (b)]. Zinc and sodium are a typical example of the first and second kind of metals, respectively.

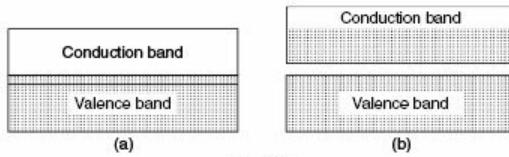


Fig. 6.2

Now, let us try to understand why these metals conduct electricity. Suppose, an electric field E is applied to the metal. Each electron in the metal experiences a force eE , gets accelerated and starts moving towards the positive electrode, i.e., in a direction opposite to that of the applied field. If l is the mean free path, then each electron gains energy equal to eEl , which works out to be of the order of the 10^{-8} eV for a typical applied field. Thus, if unoccupied states within 10^{-8} eV are available, the electrons jump into them which is the case in the metals as unoccupied states are close, i.e., within 10^{-8} eV. In such a case, the process of conduction, namely acceleration, collision and return to equilibrium, leading to steady drift of electrons, can occur. That is how, the metals conduct electricity.

- (ii) **Insulators:** The energy band structure of an insulator is shown in Fig. 6.3. The magnitude of forbidden energy gap E_g is quite large, generally 6 to 9 eV. For diamond, an insulator, it is 6 eV. This means that an electron requires a minimum of 6 eV of energy to jump from valence band to conduction band. When an electric field is applied, the electrons do not acquire sufficient energy to do so and the conduction band continues to be almost empty. So conduction is not possible and therefore insulators are very poor conductors of electricity. However, at any non-zero temperature, a few electrons (a fraction $p \propto e^{-E_g/kT}$ according to Boltzmann law) acquire sufficient energy to jump to the conduction band and therefore the conduction band is not fully empty and a small conductivity (proportional to p) results.

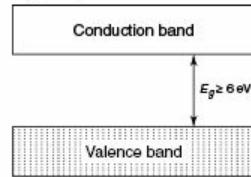


Fig. 6.3

- (iii) **Semiconductors:** The energy band structure of a semiconductor is shown in Fig. 6.4. It is similar to that of an insulator but with a smaller magnitude of forbidden energy gap, nearly 1 eV. At absolute zero temperature (0 K) the Fermi levels lie in the middle of forbidden energy gap and therefore the valence band is completely filled and the conduction band is totally empty. The absence of electrons in the conduction band at absolute zero will not allow current to flow under the influence of an applied

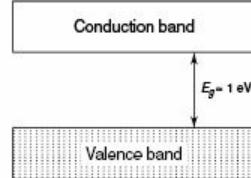


Fig. 6.4

electric field. Therefore, at absolute zero there is no difference between an insulator and a semiconductor from electrical conductivity point of view.

However, at room temperature some valence electrons acquire enough thermal energy (greater than E_g) to cross over to the conduction band and they are free to move under the influence of applied electric field. This fraction is $p \propto e^{-E_g/kT}$. Since E_g is smaller in case of semiconductors (≈ 1 eV) than that in insulators (≈ 6 eV), this fraction is sizeable for semiconductors. Thus, the crystal becomes slightly conducting at room temperature and that is why such crystals are known as semiconductors.

The forbidden energy gap is weakly dependent on temperature.

It has been found that for

$$Si : E_g(T) = 1.21 - 3.60 \times 10^{-4} T; \text{ at } T = 300 \text{ K}, E_g = 1.1 \text{ eV}$$

$$Ge : E_g(T) = 0.785 - 2.23 \times 10^{-4} T; \text{ at } T = 300 \text{ K}, E_g = 0.72 \text{ eV}$$

Thus, it decreases slightly with increasing temperature.

Higher the temperature, larger is the number of electrons jumping to the conduction band and hence greater is the conductivity. Each electron moving to the conduction band leaves behind a vacancy called a hole in the valence band, which behaves as a short of positive charge carrier.

The distinction between insulators and semiconductors is quantitative, not qualitative. In both types of solids, the valence band and the conduction band are separated by forbidden energy gap. If the forbidden energy gap $E_g \approx 1$ eV or so, the solid is categorized as semiconductor while if $E_g > 5$ eV, we say that the solid is an insulator. The carrier density (charge carriers per unit volume) and hence the conductivity depends exponentially on the ratio of the energy gap E_g to the temperature of the crystal. This ratio varies from nearly 30 for a semiconductor to about 150 for an insulator change. The typical room temperature resistivities for metals, semiconductors and insulators are $10^{-8} \Omega\text{m}$, $10^2 \Omega\text{m}$ and $10^{10} \Omega\text{m}$ respectively.

Example 6.1

The mean free path of conduction electrons in copper is about 4×10^{-8} m. In this conductor, what electrical field should be applied so that a conduction electron acquires 1.2 eV average energy?

Solution.

Let the electric field required be E . The force on an electron is eE . When an electron moves through a distance l in the field, the work done on it is eEl , which is also the energy acquired by it. An electron travels, on an average, a distance 4×10^{-8} m before collision. Therefore, average energy is $eE \times (4 \times 10^{-8}$ m). Thus,

$$eE \times (4 \times 10^{-8} \text{ m}) = 1.2 \text{ eV}$$

or

$$eE \times (4 \times 10^{-8} \text{ m}) = 1.2 \times 1.6 \times 10^{-19} \text{ J}$$

or

$$E = \frac{1.2 \times 1.6 \times 10^{-19}}{e \times 4 \times 10^{-8}} \frac{\text{V}}{\text{m}} = \frac{1.2 \times 1.6 \times 10^{-19}}{1.6 \times 10^{-19} \times 4 \times 10^{-8}} \frac{\text{V}}{\text{m}}$$

or

$$E = 3.0 \times 10^7 \text{ V/m.}$$

6.3 CURRENT CARRIERS (ELECTRONS AND HOLES) IN SEMICONDUCTORS

Figure 6.5 shows the energy band diagram of a pure semiconductor. Figure 6.5(a) shows a few electrons in the valence band at absolute zero of temperature, while Fig. 6.5(b) depicts the situation at a higher (say room) temperature. As we have discussed earlier, at higher temperatures, some of the electrons in the valence band acquire sufficient thermal energy to cross over to the conduction band. *For each electron that leaves the valence band a vacancy is created at the location where the electron was present before moving to the conduction band [Fig. 6.5(b)]. This vacancy is called a hole.*

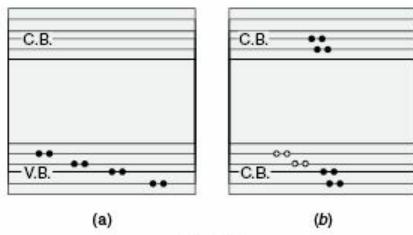


Fig. 6.5

The mechanism of hole creation can also be understood with the help of Fig. 6.6. In the figure the crystal structure of germanium, a semiconductor, is shown. The atoms in the structure are strongly held by covalent bonds. On receiving an addition energy, one of the electrons contributing to a covalent bond breaks and is free to move in the crystal. While coming out of a covalent bond, it leaves behind a hole.

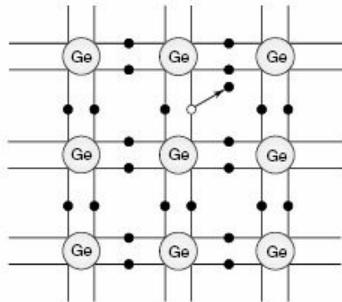


Fig. 6.6 Breaking of a covalent bond in a Ge crystal.

An electron from the neighbouring atom can break away and occupy the existing vacancy (or hole) completing the covalent bond and creating a hole at another site. In our two-dimensional example shown in Fig. 6.6, an electron from any of the neighbouring atom can come to complete the bond and the hole can move to any of these atoms. In this manner, holes move randomly in the crystal lattice. In real situations, free electrons may also fill up the vacancy. This is also known as *electron-hole recombination*.

In a crystal, the breaking of bonds or generation of electron-hole pairs and completion of bonds due to recombination occur all the time. In equilibrium, the rate of (electron-hole pair) generation equals that of recombination, maintaining a fixed number of electrons and holes.

Movement of holes under the influence of an applied electric field. Now we will try to understand that how the holes move under the influence of an applied electric field with the help of an over simplified diagram shown in Fig. 6.7. In the figure five atoms at four instants of time have been shown in a semiconductor crystal. Suppose, initially there are no free electrons and all the electrons are bound in the covalent bonds. Further suppose, that thermal excitation breaks a covalent bond at atom-1 setting an electron free and creating a hole, i.e., generates an electron-hole pair. The electron generated moves to positive terminal of the battery while the hole remains at atom-1. Due to the applied field, an electron from atom-2 goes atom-1 and fill the vacancy in atom-1 but in the process creates hole at atom-2. Similarly, an electron from atom-3 moves at atom-2 and the hole moves to atom-3. Continuing in this way, the hole moves to atom-5 which takes an electron from the negative terminal of the battery and the covalent bond is completed. Thus, the electron initially lost to the positive terminal of the battery has been finally compensated by the negative terminal of the battery and this manner an electron has travelled from negative to the positive terminal of the battery through the crystal, constituting a current in the crystal.

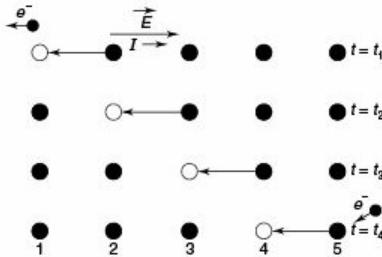


Fig. 6.7 Movement of a hole under an electric field at different instants of time in a semiconductor.

From the above discussions, we notice that holes move in a direction opposite to that of electrons, i.e., they behave as *positively charged* particles. A hole is a convenient way of describing charge motion in semiconductors, although motion can be described entirely in terms of electrons.

In a semiconductor, when both electrons and holes are present, and are, far away from each other, they carry current independently. If I_e is the component of current carried by electrons and I_h that by holes, then the total current I is given by

$$I = I_e + I_h \quad \dots(6.1)$$

Sometimes I_n for I_e for I_p and I_h are also used, in such notations the total current is $I = I_n + I_p$.

To create a hole an electron has to be removed. To remove an electron from the inner orbit, larger amount of energy is required. Thus, we can say that the hole farther from the top of the valence band has higher energy, just as an electron in the conduction band, farther away from its bottom has higher energy.

6.4 INTRINSIC SEMICONDUCTORS

Pure semiconductors are called *intrinsic semiconductors*. In intrinsic semiconductors, thermally generated electron-hole pairs exists and in the absence of any applied electric field, these electrons and holes move randomly. These electrons and holes are called *intrinsic carriers*. Because electrons and holes are created in pairs, i.e., a hole is created when an electron becomes free, in an intrinsic semiconductors

$$n = p = n_i \quad \dots(6.2)$$

where n is the electron density in conduction band, p is the hole density in the valence band and n_i is the intrinsic concentration.

In Table 6.1, energy gap (E_g), mobility (μ), intrinsic carrier concentration (n_i), conductivity (σ), resistivity, density and concentration of atom of pure Si and Ge, two widely used semiconductors, are given.

Table 6.1

	Si	Ge
Energy gap E_g (eV)	1.1	0.7
Intrinsic conductivity σ (Sm^{-1})	4.4×10^{-4}	2.18
Intrinsic resistivity ρ ($\Omega \text{ m}$)	2300	0.46
Intrinsic carrier concentration n_i (m^{-3})	1.5×10^{16}	2.4×10^{19}
Density (kgm^{-3})	2.3×10^9	5.32×10^9
Concentration of atom (m^{-3})	5×10^{28}	4.41×10^{28}
Electron mobility μ_n ($\text{m}^2 \text{ V}^{-1} \text{ s}^{-1}$)	0.135	0.39
Hole mobility μ_p ($\text{m}^2 \text{ V}^{-1} \text{ s}^{-1}$)	0.048	0.19

6.5 DOPING OF SEMICONDUCTORS: EXTRINSIC SEMICONDUCTORS

The intrinsic semiconductors, at room temperature possess thermally generated free electrons and holes but their number is too small and so that the conductivity is too small to be of any practical use. To increase their conductivity suitable impurities are added in pure semiconductors. This deliberate addition of a desirable impurity is called 'doping' and the impure atoms added are called 'dopants'. A doped semiconductor is called an extrinsic semiconductor.

The dopant atoms should be such that they preferably substitute semiconductor atoms and should not distort the crystal lattice which is possible if the dopant atoms are of same size as that of the pure semiconductor crystal atoms. The concentration of dopant atoms are generally kept below 1% of the semiconductor crystal atoms. Doping is done in a number of ways.

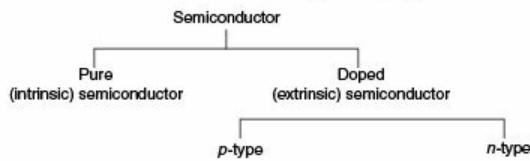
Three commonly employed methods are mentioned below:

- (i) One way is to add the impurity atoms in the melt of the semiconductor.
- (ii) Another way is to heat the crystalline semiconductor in the atmosphere containing dopant atoms or molecules, i.e., to diffuse dopant atoms into the crystalline semiconductor.
- (iii) The semiconductor to be doped is bombarded by the ions of the dopant atoms.

6.6 TYPES OF EXTRINSIC SEMICONDUCTORS: N- AND P-TYPE SEMICONDUCTORS

The most commonly used semiconductor materials used for manufacturing semiconductor materials for devices are tetravalent elements such as silicon (Si) and germanium (Ge). However, for specialized

application GaAs, SiC, etc. are becoming very useful. To increase the conductivity the pure semiconductors are doped by suitable impurities. Depending upon the kind of impurities added, there are two types of extrinsic semiconductor or : *n*-type and *p*-type.



N-Type (Extrinsic) Semiconductor. If we dope intrinsic semiconductor, say silicon, which has four valence electrons, with a controlled amount of pentavalent atoms like arsenic As (or antimony Sb or phosphorus P), which have five valence electrons, the added atoms substitute the silicon atoms [Fig. 6.8(a)]. Out of the five valence electrons of an arsenic atom, four electrons enter in covalent bonds with the four silicon atoms surrounding it, while the fifth electron not being part of any bond, is comparatively free. Obviously, the number of free electrons will be equal to the number of atoms added. Thus, *each pentavalent atom added releases one free electron to the host crystal and it is for this reason that the pentavalent atoms are called the donor atoms. Because of these free electrons in the doped crystal, the doped crystal is called n-type Si crystal*. The electrons set free are called *extrinsic carriers* and the semiconductor is called *n*-type extrinsic semiconductor. On giving up the fifth electron the donor atom acquires positive charge although the crystal as a whole remains electrically neutral.

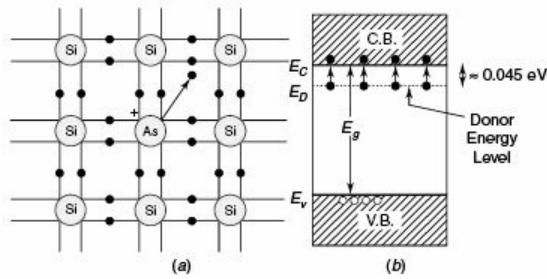


Fig. 6.8 N-type semiconductor.

Like pure semiconductors, thermally generated electron-hole pairs exist in doped semiconductors also. But due to excess (free) electrons in the *n*-type extrinsic semiconductors, *the number of free electrons is much higher than that of holes. For this reason, in n-type semiconductors, electrons are called majority carriers and holes are called minority carriers*.

The fifth electron of the donor atom, not a part of any covalent bond, is very loosely bound to the nucleus. This we can realize from the following discussion. The electron moves in crystalline silicon (or germanium) which has a relative dielectric constant of nearly ten. This means the coulomb attraction between the electron and the nucleus is ten times smaller than what it would have been in

free space. The binding energy of this electron works out to be only 0.45 eV. Thus, to free the fifth electron from the donor atom only 0.045 eV of energy is needed. In terms of band structure, it is equivalent to say that *such electrons (extrinsic carriers) create a donor energy level just below (0.045 eV) the conduction bond*, as shown in Fig. 6.8(b). This energy (0.045 eV) is comparable to room temperature thermal energy $kT \approx 0.03$ eV, and is much smaller than the energy gap of 1.1 eV.

Since it requires only 0.045 eV of energy by these electrons to go to the conduction band, at room temperature, a large fraction of donor electrons is in the conduction band giving a good conductivity. Even for low donor concentration the number of extrinsic carriers are much larger than that of intrinsic carriers at non-zero temperatures. At room temperature, most of the donor atoms are ionized, so that to a good approximation we can assume that the fifth electron from all the donor atoms are in the conduction band.

In a doped semiconductor, the number densities of conduction band electrons n and the valence band holes p differ from that in a pure semiconductor. If n_i is the number density of intrinsic carriers, then we have the relation.

$$np = n_i^2 \quad \dots(6.3)$$

In *n*-type semiconductors, the number density of electrons in conduction band is much larger than that of holes in valence band and is nearly equal the density of donor atoms (N_D) in the doped semiconductor, that is

$$n_e = N_D \gg p \quad \dots(6.4)$$

In other words, electrons are majority carriers and holes are minority carriers.

P-Type (Extrinsic) Semiconductor. If we dope an intrinsic semiconductor, say silicon, with a controlled amount of trivalent atoms, say indium (In) or boron B or aluminium Al which have three valence electrons, the added atoms substitute some Si atoms [Fig. 6.9(a)]. *Because the dopant atom has only three valence electrons, there is one incomplete bond with a neighbouring Si atom, due to deficiency of an electron. The bond is completed by taking an electron from one of the neighbouring Si-Si bonds. In the process an atom acquires negative charge and a hole is created in the crystal, as shown in Fig. 6.9(b). As the trivalent dopant atoms accept electrons from the silicon crystal, they are also called acceptor atoms.* The semiconductor obtained by adding acceptor atoms is called *p*-type (extrinsic) semiconductor as it contains excess holes. Each hole is equivalent to a positive charge.

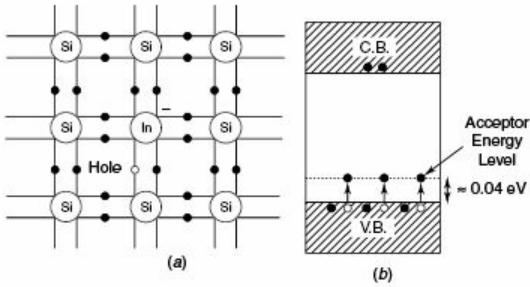


Fig. 6.9 *p*-type semiconductor.

In *p*-type semiconductors besides, thermally generated holes (due to creation of electron-hole pairs) dopant atoms create one hole per added atom. Therefore, number of holes is much larger than free electrons and that is why *holes are majority carriers and electrons are minority carriers*.

The hole produced is attracted to the negatively charged acceptor nucleus. Such holes create an acceptor energy level just above (≈ 0.04 eV) the top of the valence band [Fig. 6.9(b)]. This means that electrons in the valence band require only nearly 0.04 eV of energy to raise themselves to acceptor energy level, which is comparable to the room temperature thermal energy, and thus create holes in the valence band.

Number density of holes p in the valence band in a *p*-type semiconductor is approximately equal to the number density of added acceptor atoms N_A and is very large compared to the number density of electrons n . That is,

$$p \approx N_A \gg n_e \quad \dots(6.5)$$

6.7 COMPARISON BETWEEN INTRINSIC AND EXTRINSIC SEMICONDUCTOR

<i>Intrinsic (Pure) Semiconductor</i>	<i>Extrinsic (Doped) Semiconductor</i>
1. Intrinsic semiconductor does not contain any added impurities.	Extrinsic semiconductor contains suitable impurities added to it.
2. Its electrical conductivity is very low.	Its electrical conductivity is high.
3. There is no classification of intrinsic semiconductors.	Extrinsic semiconductors are of two types: <i>n</i> -type and <i>p</i> -type.
4. The number of free electrons in conduction band is always equal to the number of holes in valence band.	In <i>n</i> -type semiconductors, the number of electrons in conduction band is larger than that of holes in valence band while reverse is true in a <i>p</i> -type semiconductor.
5. Rarely used in semiconductor devices.	Widely used in semiconductor devices.
6. Its electrical conductivity is a function of temperature only.	Its electrical conductivity is dependent on the amount of doping as well as temperature.
7. Examples are pure Si and Ge crystals.	Examples are arsenic (or phosphorus or antimony) doped Si (or Ge) crystal for <i>n</i> -type and indium (or Al or boron) doped Si (or Ge) crystal for <i>p</i> -type.

6.8 COMPARISON BETWEEN N- AND P-TYPE SEMICONDUCTORS

<i>N-Type Semiconductor</i>	<i>P-Type Semiconductor</i>
1. N-type semiconductor is obtained by doping intrinsic semiconductor by pentavalent impurities such as arsenic or antimony or phosphorus.	P-type semiconductor is obtained by doping intrinsic semiconductor by trivalent impurities such as indium, aluminum or boron.
2. Electrons are majority carriers and holes are minority carriers.	Holes are majority carriers and electrons are minority carriers.
3. The added impurity atoms are called donor atoms as they donate one excess electron per added atom.	The added impurity atoms are called acceptor atoms as they accept one electron per added atom.

Contd..

Contd..

- | | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4. Donor energy level is just below the conduction band.
5. Due to higher mobility of electrons, for the same level of dopings as in a p-type semiconductor, it has higher electrical conductivity. | Acceptor energy level is just above the valence band.
Due to comparatively lower mobility of holes for the same level of doping as in a n-type semiconductor, it has lower electrical conductivity. |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

6.9 CHARACTERISTICS OF A SEMICONDUCTOR

The important characteristics of semiconductors are listed below.

1. At absolute zero and very low temperature a pure semiconductor behaves as an insulator. As temperature is increased few electrons in the valence band acquire energy greater than forbidden gap energy and move to conduction band. At room temperature it becomes slightly conducting.
2. The conductivity of semiconductors increases with increasing temperature. This is in contrast with the behaviour of metals in which conductivity decreases with increasing temperature.
3. There are two types of carriers in semiconductors, namely electrons and holes.
4. Very small amount of doping (≈ 1 impurity atom in 10^6 atoms) improves the conductivity drastically.
5. Doped semiconductors do not obey Ohm's law.
6. In semiconductors electrons have higher mobility than that of holes.

6.10 ELECTRICAL CONDUCTIVITY OF EXTRINSIC SEMICONDUCTORS

Consider a slab of an extrinsic semiconductor having length l and area of cross-section A .

Let n be the electron density (number of electrons per unit volume) and p that of holes. When a potential difference V is applied across it, a current I flows through it, as shown in Fig. 6.10. In a semiconductor the total current I consists of two components, the electronic current I_e due to motion of electrons (in conduction band) directed towards positive terminal and current I_p due to motion of holes (in valence band) directed towards negative terminal. Thus,

$$I = I_e + I_p \quad \dots(6.6)$$

We know the current I in a conductor having area of cross-section A is given by

$$I = enA v_d$$

where n is the number density of charge carrier, e is the charge on a carrier and v_d , the drift velocity of the carriers. For the electronic current, we have

$$I_e = enA v_e \quad \dots(6.7)$$

where n is the number density of electrons, e the electronic charge and v_e is drift velocity of electrons.

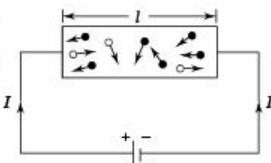


Fig. 6.10

$$I_p = e_p A v_h \quad \dots(6.8)$$

where p is number density and v_h drift velocity of holes. The charge e on a hole is equal in magnitude to that on an electron but opposite, i.e., positive in sign.

From Eqs. (6.6), (6.7) and (6.8), we have

$$\begin{aligned} I &= en A v_e + ep A v_h \\ \text{or} \quad I &= e A (n v_e + p v_h) \end{aligned} \quad \dots(6.9)$$

Using the relation $V = IR$ or $I = \frac{V}{R}$, we have

$$\frac{V}{R} = e A (n v_e + p v_h) \quad \dots(6.10)$$

Also, $R = \rho \frac{l}{A}$

where ρ is the resistivity of the semiconductor. $\dots(6.11)$

From Eqs. (6.10) and (6.11), we get

$$\begin{aligned} \frac{VA}{\rho l} &= e A (n v_e + p v_h) \\ \text{or} \quad \frac{V}{\rho l} &= e (n v_e + p v_h) \end{aligned}$$

But $\frac{V}{l} = E$ is the electric field set up across the semiconductor.

$$\therefore \frac{E}{\rho} = e (n v_e + p v_h) \quad \dots(6.12)$$

Dividing the above equation by E , we have

$$\begin{aligned} \frac{1}{\rho} &= \frac{e}{E} (n v_e + p v_h) \\ \text{or} \quad \frac{1}{\rho} &= e \left(n \frac{v_e}{E} + p \frac{v_h}{E} \right) \end{aligned}$$

The quantities $\frac{v_e}{E} = \mu_e$ and $\frac{v_h}{E} = \mu_h$ denote mobility of electrons and holes respectively. Thus,

$$\begin{aligned} \frac{1}{\rho} &= e (n \mu_e + p \mu_h) \\ \text{or, conductivity} \quad \sigma &= \frac{1}{\rho} = e (n \mu_e + p \mu_h) \end{aligned} \quad \dots(6.13)$$

From Eq. (6.13) it is clear that *conductivity is proportional to electron and hole densities and their mobilities*. The mobility of electrons and holes are weakly dependent on temperature. The

increase in conductivity with increasing temperature takes mainly due to the increase in n and p . The mobility of electron is slightly higher than that of holes.

Example 6.2

Pure Si at 300 K has equal electron (n_i) and hole (p) concentration of $1.5 \times 10^{16} \text{ m}^{-3}$. Doping by indium increases p to $4.5 \times 10^{22} \text{ m}^{-3}$. Calculate n in the doped silicon.

Solution.

Here, $n_i = 1.5 \times 10^{16} \text{ m}^{-3}$
 $p = 4.5 \times 10^{22} \text{ m}^{-3}$

We have the relation $np = n_i^2$

$$\therefore n = \frac{n_i^2}{p} = \frac{1.5 \times 10^{16} \times 1.5 \times 10^{16}}{4.5 \times 10^{22}} \text{ m}^{-3}$$

$$= 5 \times 10^9 \text{ m}^{-3}.$$

Example 6.3

For an n -type semiconducting material following data is given: density of electron = $8 \times 10^{22} \text{ cm}^{-3}$, density of holes $5 \times 10^{12} \text{ cm}^{-3}$ mobility of conduction electron = $2.3 \times 10^4 \frac{\text{cm}^2}{\text{V} - \text{s}}$ and mobility of holes = $100 \text{ cm}^2/\text{V} - \text{s}$. Calculate the resistivity of the material.

Solution.

We know that the conductivity of a semiconductor is

$$\sigma = e(n\mu_e + p\mu_h)$$

Here, $n = 8 \times 10^{13} \text{ cm}^{-3} = 8 \times 10^{19} \text{ m}^{-3}$, $\mu_e = 2.3 \times 10^4 \frac{\text{cm}^2}{\text{V} - \text{s}} = 2.3 \frac{\text{m}^2}{\text{V} - \text{s}}$

$$p = 5 \times 10^{12} \text{ cm}^{-3} = 5 \times 10^{18} \text{ m}^{-3}$$
, $\mu_h = 100 \text{ cm}^2/\text{V} - \text{s} = 10^{-2} \text{ m}^2/\text{V} - \text{s}$

and $e = 1.6 \times 10^{-19} \text{ C}$

$$\therefore \sigma = 1.6 \times 10^{-19} (8 \times 10^{19} \times 2.3 + 5 \times 10^{18} \times 10^{-2}) \frac{\text{C}}{\text{mVs}}$$

$$= 2.94 \frac{\text{C}}{\text{mVs}}$$

Resistivity $\rho = \frac{1}{\sigma} = \frac{1}{2.94} \Omega - \text{m} = 0.34 \Omega - \text{m}$

Example 6.4

Find the resistivity of intrinsic silicon using the following data: doped with pentavalent impurity atom for each 60 million silicon atoms. What is the new resistivity? Given, density of Si atom = 4.5

$\times 10^{28} \text{ m}^{-3}$, intrinsic carrier concentration = $1.5 \times 10^{16} \text{ m}^{-3}$, electron mobility = $0.135 \frac{\text{m}^2}{\text{Vs}}$ and hole

mobility = $0.048 \frac{\text{m}^2}{\text{Vs}}$.

Solution.

The conductivity of the intrinsic silicon

$$\sigma_i = e n_i (\mu_e + \mu_h) \quad (\because n_e = n_h = n_i)$$

$$\text{Here, } n_i = 1.5 \times 10^{16} \text{ m}^{-3}, \mu_e = 0.135 \frac{\text{m}^2}{\text{Vs}}, \mu_h = 0.048 \frac{\text{m}^2}{\text{Vs}}$$

$$\begin{aligned} \therefore \sigma_i &= 1.6 \times 10^{-19} \times 1.5 \times 10^{16} (0.135 + 0.048) \frac{\text{mho}}{\text{m}} \\ &= 2.4 \times 10^{-3} \times 0.183 \text{ mho/m} \\ &= 0.4392 \times 10^{-3} \text{ mho/m} \end{aligned}$$

$$\begin{aligned} \text{Thus, intrinsic resistivity } \rho_i &= \frac{1}{\sigma_i} = \frac{1}{0.4392} \times 10^{-3} \text{ ohm-m} \\ &= 2.27 \times 10^3 \text{ ohm-m} \end{aligned}$$

Concentration of donor atoms

$$\begin{aligned} N_D &= \frac{\text{Concentration of Si atom (N)}_i}{60 \times 10^6} \\ &= \frac{4.5 \times 10^{28}}{6.0 \times 10^7} \text{ m}^{-3} = 7.5 \times 10^{20} \text{ m}^{-3} \end{aligned}$$

Hence, concentration of free electrons

$$n = N_D = 7.5 \times 10^{20} \text{ m}^{-3}$$

\therefore Hole concentration (density)

$$\begin{aligned} p &= \frac{n_i^2}{n} = \frac{1.5 \times 10^{16} \times 1.5 \times 10^{16}}{7.5 \times 10^{20}} \text{ m}^{-3} \\ &= 3.0 \times 10^{11} \text{ m}^{-3} \end{aligned}$$

Thus, we see that $n_e \gg p$, i.e., the semiconductor is *n*-type. The new conductivity of the semiconductor becomes

$$\sigma = e (n \mu_e + p \mu_h)$$

$$\begin{aligned}
 &= 1.6 \times 10^{-19} (7.5 \times 10^{20} \times 0.135 + 3.0 \times 10^{11} \times 0.048) \frac{\text{mho}}{\text{m}} \\
 &= 1.6 \times 10^{-19} (7.5 \times 10^{20} \times 0.135) \frac{\text{mho}}{\text{m}} \\
 &\quad (\text{Contribution of second term is negligible}) \\
 &= 0.162 \text{ mho/m}
 \end{aligned}$$

\therefore Resistivity of doped semiconductor

$$\rho = \frac{1}{\sigma} = \frac{1}{0.162} = 6.172 \Omega \text{ m.}$$

Example 6.5

In a semiconductor it is observed that three-quarters of the current is carried by electrons and one quarter by holes. If at this temperature, the drift speed of electrons is three times that of the holes, determine the ratio of electrons to holes in the semiconductors.

Solution.

In a semiconductor, total current $I = I_e + I_h$

$$\text{Here, } I_e = \frac{3}{4} I \text{ and } I_h = \frac{1}{4} I$$

also,

$$v_e = 3 v_h$$

$$I_e = e n A v_e \text{ and } I_h = e p A v_h$$

$$\text{Here } \frac{I_e}{I_h} = \frac{e n A v_e}{e p A v_h} = \frac{n v_e}{p v_h} = \frac{n 3 v_h}{p v_h} = \frac{3n}{p}$$

$$\therefore \frac{n}{p} = \frac{1}{3} \cdot \frac{I_e}{I_h} = \frac{1}{3} \cdot \frac{\frac{3}{4} I}{\frac{1}{4} I} = 1$$

$$\Rightarrow n = p$$

Thus, the ratio of density of electrons to that of holes present in the semiconductor is unity.

Example 6.6

Find the resistivity of an intrinsic silicon doped with the donor impurity to the extent of 1 in 10^8 atoms. Given: density of silicon atoms = $5 \times 10^{28} \text{ m}^{-3}$, $m_e = 0.135$, $n_i = 1.5 \times 10^{16} \text{ m}^{-3}$.

Solution.

The density of donor atoms

$$\begin{aligned}
 N_D &= \frac{5 \times 10^{28}}{10^8} = 5 \times 10^{20} \text{ m}^{-3} \\
 np &= n_i^2 = (1.5 \times 10^{16})^2 = 2.25 \times 10^{32} \text{ m}^{-3}
 \end{aligned}$$

Since $n = N_D$,

$$\therefore p = \frac{n_i^2}{N_D} = \frac{2.25 \times 10^{32}}{5 \times 10^{20}} = 4.5 \times 10^{11} \text{ m}^{-3}$$

Since $n (= 5 \times 10^{20}) \gg p (= 4.5 \times 10^{11})$, clearly electrons are majority carriers and the material is *n*-type. The conductivity $\sigma = e n_e \mu_e = 5 \times 10^{20} \times 1.6 \times 10^{-19} \times 0.135 \text{ mho/m} = 10.8 \text{ mho/m}$.

$$\therefore \text{Resistivity, } \rho = \frac{1}{\sigma} = \frac{1}{10.8} \Omega\text{m} = 0.0926 \Omega\text{m}$$

Thus, we see that resistivity drops drastically to $0.0926 \Omega\text{m}$ from that of intrinsic semiconductor $2300 \Omega\text{m}$.

6.11 CARRIER CONCENTRATION IN INTRINSIC SEMICONDUCTORS

In intrinsic semiconductors, electrons in the conduction band and holes in the valence band constitute the charge carriers. Since these charge carriers are generated due to the breaking of covalent bonds, in the intrinsic semiconductors, the number of electrons is equal to the number of holes. At 0 K, since the covalent bonds are intact, the semiconductors act as insulators since hardly any charge carriers are present. As the temperature increases, the covalent bonds are broken and electron-hole pairs are created. We obtain below the expressions for the number of electrons in the conduction band per unit volume (n) and the number of holes in the valence band per unit volume (p).

6.11.1 Density of Electrons

Let dn be the number of electrons in the energy interval E and $E + dE$ in the conduction band,

$$dn = Z(E) F(E) dE \quad \dots(6.14)$$

where $Z(E) dE$ is the density of states in the energy interval E and $E + dE$ and $F(E)$ is the electron occupancy probability, i.e., the probability of a state of energy E being occupied by an electron.

If E_C is the energy corresponding to the bottom of the conduction band, to obtain the number of electrons in the conduction band, Eq. (6.14) has to be integrated from E_C to the top of the conduction band,

$$n = \int_{E_C}^{\infty} Z(E) F(E) dE \quad \dots(6.15)$$

The upper limit of integration is taken as infinity since the probability of electrons occupying the upper levels of the conduction band reduces to zero at infinity.

The density of states, i.e., the number of energy states per unit volume within the energy interval E and $E + dE$ is given by

$$Z(E) dE = \frac{4\pi}{h^3} (2m)^{3/2} E^{1/2} dE \quad \dots(6.16)$$

To account for the possibility of two electrons of opposite spin occupying any energy state, corresponding to Pauli's exclusion principle, the above expression has been multiplied by 2 compared to the original expression. Since the electron moves in a periodic potential, its mass has to be

replaced by its effective mass m_e^* . Hence, the density of electrons in the conduction band is given by

$$Z(E) dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} E^{1/2} dE$$

Since the E starts at the bottom of the conduction band E_C ,

$$Z(E) dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_C)^{1/2} dE \quad \dots(6.17)$$

We know the probability of an electron occupying an energy state E is given

$$F(E) dE = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \quad \dots(6.18)$$

where E_F is the energy of the Fermi level which is exactly at the middle of the forbidden energy gap in the intrinsic semiconductor, i.e.,

$$E_F = \frac{E_C + E_V}{2}$$

where, E_C is the energy of the bottom of the conduction band and E_V is the energy of the top of the valence band. Equation (6.18) can also be written as

$$F(E) dE = \left[1 + \exp\left(\frac{E - E_F}{kT}\right) \right]^{-1}$$

For all possible temperatures, $E - E_F \gg kT$,

$$\therefore F(E) dE \approx \exp\left(\frac{E_F - E}{kT}\right)$$

Therefore, Eq. (6.15) becomes

$$\begin{aligned} n &= \int_{E_C}^{\infty} \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_C)^{1/2} \exp\left(\frac{E_F - E}{kT}\right) dE \\ &= \frac{4\pi}{h^3} (2m_e^*)^{3/2} \int_{E_C}^{\infty} (E - E_C)^{1/2} \exp\left(\frac{E_F - E}{kT}\right) dE \\ &= \frac{4\pi}{h^3} (2m_e^*)^{3/2} \exp\left(\frac{E_F}{kT}\right) \int_{E_C}^{\infty} (E - E_C)^{1/2} \exp\left(-\frac{E}{kT}\right) dE \quad \dots(6.19) \end{aligned}$$

To evaluate this integral, let $E - E_C = x$

$$\therefore E = E_C + x \Rightarrow dE = dx$$

Substituting these, we get

$$n = \frac{4\pi}{h^3} (2m_e^*) \exp\left(\frac{E_F}{kT}\right) \int_0^{\infty} x^{1/2} \exp\left(-\frac{(E_C + x)}{kT}\right) dx$$

$$= \frac{4\pi}{h^3} (2m_e^*)^{3/2} \exp\left(\frac{(E_F + E_C)}{kT}\right) \int_0^\infty x^{1/2} \exp\left(-\frac{x}{kT}\right) dx \quad \dots(6.20)$$

Using gamma function it can be shown that

$$\int_0^\infty x^{1/2} \exp\left(-\frac{x}{kT}\right) dx = (kT)^{3/2} \frac{\pi^{1/2}}{2}$$

Hence,

$$n = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \left(\exp\left(\frac{E_F - E_C}{kT}\right) (kT)^{3/2} \frac{\pi^{1/2}}{2} \right)$$

Thus, the **number of electrons per unit volume (or concentration of electron) in the conduction band** is given by

$$n = 2 \left(\frac{2\pi m_e^* kT}{h^2} \right)^{3/2} \exp\left(\frac{E_F - E_C}{2}\right) \quad \dots(6.21)$$

6.11.2 Density of Holes

Let dp be the number of holes in the valence band within the energy interval E and dE , given by

$$dp = Z(E) (1 - F(E)) dE \quad \dots(6.22)$$

where $Z(E) dE$ is the density of states in the energy interval E and $E + dE$ and $(1 - F(E))$ is the probability of existence of a hole. Here, it is important to note that since $F(E)$ is the probability of electron occupancy, $1 - F(E)$ gives the probability of the absence of an electron or the presence of a hole. Thus,

$$\begin{aligned} 1 - F(E) &= 1 - \left[\frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \right] \\ &= 1 - \left[1 + \left(\exp\left(\frac{E - E_F}{kT}\right) \right) \right]^{-1} \end{aligned}$$

Since $E - E_F \gg kT$,

$$\therefore 1 - F(E) = \exp\left(\frac{E - E_F}{kT}\right) \quad \dots(6.23)$$

Density of holes in the valence band

$$Z(E) dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} E^{1/2} dE$$

Since E_V is the energy of the top of the valence band

$$Z(E) dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_V - E)^{1/2} dE$$

To calculate the number of holes in the valence band, the above equation has to be integrated from $-\infty$ to the energy corresponding to the top of the conduction band E_V (for mathematical simplicity, the lowest energy is taken as $-\infty$). Thus, the number of holes in the valence band per unit volume is given by

$$\begin{aligned} p &= \int_{-\infty}^{E_V} \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_V - E)^{1/2} \exp\left(\frac{E - E_F}{kT}\right) dE \\ &= \frac{4\pi}{h^3} (2m_h^*) \exp\left(-\frac{E_F}{kT}\right) \int_{-\infty}^{E_V} (E_V - E)^{1/2} \exp\left(\frac{E}{kT}\right) dE \end{aligned}$$

To evaluate this integral, let $E_V - E = x$

$$\text{or } E = E_V - x \Rightarrow dE = -dx$$

$$\begin{aligned} \therefore \int_{-\infty}^{E_V} (E_V - E)^{1/2} \exp\left(\frac{E}{kT}\right) dE &= - \int_{\infty}^0 x^{1/2} \exp\left(\frac{E_V - x}{kT}\right) dx \\ &= \exp\left(\frac{E_V}{kT}\right) \int_0^{\infty} x^{1/2} \exp\left(-\frac{x}{kT}\right) dx \\ &= \exp\left(\frac{E_V}{kT}\right) (kT)^{3/2} \frac{\pi^{1/2}}{2} \quad (\text{using gamma } f_n) \end{aligned}$$

$$\text{Hence, } p = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp\left(\frac{E_V - E_F}{kT}\right) (kT)^{3/2} \frac{\pi^{1/2}}{2}$$

$$\text{or } p = 2 \left(\frac{2\pi m_h^* kT}{h^2} \right)^{3/2} \exp\left(\frac{E_V - E_F}{kT}\right) \quad \dots(6.24)$$

6.11.3 Intrinsic Carrier Concentration

In intrinsic semiconductors, $n = p$.

Hence, $n = p \equiv n_i$ is called the **intrinsic carrier concentration** and is given by

$$\begin{aligned} n_i^2 &= np = 4 \left(\frac{2\pi kT}{h^2} \right)^3 (m_e^* m_h^*)^{3/2} \exp\left(\frac{E_V - E_C}{kT}\right) \\ &= 4 \left(\frac{2\pi kT}{h^2} \right)^3 (m_e^* m_h^*) \exp\left(\frac{-E_g}{kT}\right) \end{aligned}$$

where E_g is the forbidden energy gap. Hence,

$$n_i = 2 \left(\frac{2\pi kT}{h^2} \right)^{3/2} (m_e^* m_h^*)^{3/2} \exp\left(\frac{-E_g}{kT}\right) \quad \dots(6.25)$$

Fermi Level in Intrinsic Semiconductor

In intrinsic semiconductors,

$$n = p$$

$$\text{or, } 2 \left(\frac{2\pi m_e^* kT}{h^2} \right)^{3/2} \exp \left(\frac{E_F - E_C}{kT} \right) = 2 \left(\frac{2\pi m_h^* kT}{h^2} \right)^{3/2} \exp \left(\frac{E_V - E_F}{kT} \right)$$

$$\text{or, } (m_e^*)^{3/2} \exp \frac{E_F - E_C}{kT} = (m_h^*)^{3/2} \exp \frac{E_V - E_F}{kT}$$

$$\text{or, } \exp \frac{2 E_F}{kT} = \left(\frac{m_h^*}{m_e^*} \right)^{3/2} \exp \left(\frac{E_V + E_C}{kT} \right)$$

Taking Logarithms on both sides

$$\frac{2 E_F}{kT} = \frac{3}{2} \log \frac{m_h^*}{m_e^*} + \left(\frac{E_V + E_C}{kT} \right)$$

$$\text{or, } E_F = \frac{3 kT}{4} \log \frac{m_h^*}{m_e^*} + \left(\frac{E_V + E_C}{2} \right) \quad \dots(6.26)$$

Assuming $m_h^* = m_e^*$ then

$$E_F = \frac{E_V + E_C}{2} \quad \dots(6.27)$$

Thus, in an intrinsic semiconductor, Fermi level is located at the middle of the valence and conduction bands and its position is independent of temperature (Fig. 6.11). However, since $m_h^* > m_e^*$, E_F is just above the middle and rises slightly with increasing temperature.

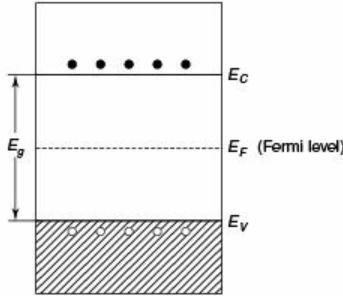


Fig. 6.11 Fermi level at the middle of valence and conduction bands.

6.12 ELECTRICAL CONDUCTIVITY: VARIATION WITH TEMPERATURE

We know that for a semiconductor, electrical conductivity σ is given by

$$\sigma = ne \mu_n + pe \mu_n$$

where μ_n and μ_h are the mobilities of electrons and holes respectively. Mobility is velocity acquired by the electron or hole per unit electric field.

For intrinsic semiconductor, $n = p \equiv n_i$

$$\therefore \sigma = n_i e (\mu_e + \mu_h)$$

Substituting for n_i from Eq. (6.25), we get

$$\sigma = (\mu_e + \mu_h) 2e \left(\frac{2\pi kT}{h^2} \right)^{3/2} (m_e^* m_h^*)^{3/4} \exp \left(-\frac{E_g}{2kT} \right) \quad \dots(6.28)$$

Thus, the electrical conductivity depends upon the negative exponential of the forbidden gap energy between the valence and conduction bands and on the mobilities of both holes and electrons. The mobilities are determined by the interaction of the electron or hole with lattice waves or phonons. In such a case μ_e and μ_h are both proportional to $T^{-3/2}$.

Hence, the electrical conductivity of an intrinsic semiconductor can be expressed as

$$\sigma_i = A \exp \left(-\frac{E_g}{2kT} \right) \quad \dots(6.29)$$

where A is a constant.

Taking logarithms on both sides, we have

$$\log \sigma_i = \log A - \frac{E_g}{2kT} \quad \dots(6.30)$$

A typical graph between $\frac{1}{T}$ and $\log \sigma_i$ is shown in Fig. 6.12. The graph indicates that the conductivity increases with temperature.

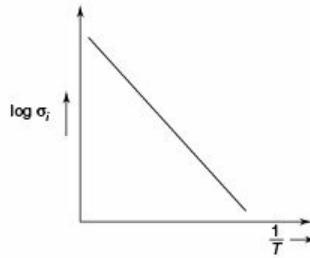


Fig. 6.12 Variation of electrical conductivity with temperature.

With increase temperature, more and more of the covalent bonds get broken and electrons and holes are generated, increasing the conductivity. Although, the carrier mobility decreases due to the increased mobility, the increased number of charge carriers overcome this and the net effect is the increase of σ with increasing temperature.

6.13 DRIFT AND DIFFUSION CURRENTS: EINSTEIN'S EQUATION

The net current that flows across a semiconductor has two components: (i) drift current and (ii) diffusion current.

6.13.1 Drift Current

In the absence of any externally applied electric field, the random motion of free carriers within a crystal does not result in a net transfer of charge since the charge transfer in one direction is balanced by movement of charge in the opposite direction. In other words, there is no flow of current. When a voltage is applied across the crystal each charge carrier experiences a force, electrons towards positive terminal and the holes towards the negative terminal. This results in a net flow of charge and the resulting current is called the *drift current*. The corresponding current densities due to drift of electrons and holes respectively are given by

$$J_n(\text{drift}) = n \mu_n e E$$

$$J_p(\text{drift}) = p \mu_p e E$$

Therefore, the total drift current density is given by

$$\begin{aligned} J(\text{drift}) &= J_n(\text{drift}) + J_p(\text{drift}) \\ &= n \mu_n e E + p \mu_p e E \end{aligned} \quad \dots(6.31)$$

Comparing this with Ohm's law $J = \sigma E$, gives

$$\sigma = n \mu_n e + p \mu_p e$$

For an intrinsic semiconductor, $n = p = n_i$. Therefore, the intrinsic conductivity is given by

$$\sigma_i = n_i e (\mu_n + \mu_p) \quad \dots(6.32)$$

For extrinsic semiconductors, only one term will be dominant as the current due to the minority carrier is negligible.

6.13.2 Diffusion Current

In addition to drift current, charge carrier in the semiconductors can also move by the *diffusion process*. The motion of the carrier by diffusion occurs when there is a non-uniform distribution of charged particles. The free electron density, at equilibrium, in a homogeneous semiconductor is uniform in the absence of any externally applied electric field. However, if excess carriers are introduced in a small region of the crystal either by injecting carriers or by generating carriers by heating or radiating, a *non-uniform distribution of charge carriers* is created. Ignoring the recombinations taking place, the *excess carriers move from the region of higher density to the regions of lower density tending to produce a uniform distribution*.

Let us assume that the concentration Δn of electrons varies in a distance Δx , in the limit as $\Delta x \rightarrow 0$. Thus, a *concentration gradient* $\frac{dn}{dx}$ of the electrons exists in the material (ignoring the time variation of n). The electron flux at any point due to diffusion is proportional to the concentration gradient at that point, i.e.,

$$\text{Diffusion electron flux} \propto \frac{dn}{dx} = D_n \frac{dn}{dx}$$

where, D_n is a proportionality constant called the *diffusing constant for electrons*. Following the Ficks's law which states the movement of charge carriers is in the direction of negative gradient. Therefore, the rate of flow electrons across unit area is equal to

$$-D_n \frac{dn}{dx}$$

This flow constitutes an *electron diffusion current density* and since conventional current is the rate of positive charge.

$$\begin{aligned} J_n(\text{diff}) &= -e \text{ (rate of flow of electrons across per unit area)} \\ &= e D_n \frac{dn}{dx} \end{aligned} \quad \dots(6.33)$$

Similarly, if excess holes are created in a region, hole diffusion takes place and constitute a hole diffusion current given by

$$\begin{aligned} J_p(\text{diff}) &= (+e) \text{ (rate of diffusion of holes across per unit area)} \\ &= (+e) \left(-D_p \frac{dp}{dx} \right) \\ &= -e D_p \frac{dp}{dx} \end{aligned} \quad \dots(6.34)$$

where D_p is the hole diffusion coefficient.

If in a semiconducting material, an electric field E as well as a concentration gradient exists in a given direction say x , then the total current due to electrons is given by

$$\begin{aligned} J_{nx} &= J_n(\text{drift}) + J_n(\text{diff}) \\ &= n_e \mu_n E + e D_n e \frac{dn}{dx} \end{aligned} \quad \dots(6.35)$$

Similarly, total current density due to holes is given by

$$\begin{aligned} J_{px} &= J_p(\text{drift}) + J_p(\text{diff}) \\ &= p e \mu_p E - e D_p \frac{dp}{dx} \end{aligned} \quad \dots(6.36)$$

Thus, total current density is given by

$$J_x = J_n + J_p = e \left[\left(n \mu_n E + D_n \frac{dn}{dx} \right) + \left(p \mu_p E - D_p \frac{dp}{dx} \right) \right] \quad \dots(6.37)$$

The mobility μ express the ability of carriers drift while the diffusion constant D expreses the ability of carriers to diffuse.

6.13.3 Einstein's Equation

Both the drift and diffusion processes are dependent on the scattering which hinders the movement of charge carriers. Einstein showed that the parameters describing drift and diffusion, mobility μ and diffusion constant D respectivitely, are directly related.

At equilibrium with no applied electric field, the free electron distribution is uniform and there is no net current flow. Any change in the state of equilibrium, which would lead to a diffusion current, creates an internal electric field and a drift current balancing the diffusion current component. Thus, in equilibrium

$$\Delta n e E \mu_n = e D_n \frac{\partial \Delta n}{\partial x} \quad \dots(6.38)$$

The force F on excess carriers restoring equilibrium is given by

$$F = (\Delta n) e E = \frac{e D_n}{\mu_n} \frac{\partial \Delta n}{\partial x} \quad \dots(6.39)$$

This force F on excess carriers is also dependent on the thermal energy of these carriers. Making an analogy between the excess carriers in a semiconductor and gas molecules in a gas container, the force corresponding to the pressure gradient is equal to $kT \frac{\partial(\Delta n)}{\partial x}$ as from the kinetic theory of gases, gas pressure is nkT , where n is the molecular concentration. Comparing this with Eq. (6.39), we get

$$kT = \frac{e D_n}{\mu_n}$$

or

$$D_n = \frac{\mu_n kT}{e} \quad \dots(6.40)$$

This relation between diffusion coefficient and mobility of a charge carrier is known as *Einstein's equation*.

Similarly, for the holes, this relation is given by

$$D_p = \frac{\mu_p kT}{e} \quad \dots(6.41)$$

Hence, from Eqs. (6.40) and (6.41), we have

$$\frac{D_n}{D_p} = \frac{\mu_e}{\mu_p} \quad \dots(6.42)$$

Thus, knowing the other values, the diffusion coefficients of the charge carriers can be calculated. The calculated value of D_e for Si is $3.4 \times 10^{-3} \text{ m}^2/\text{s}$ and that of Ge is $9.9 \times 10^{-3} \text{ m}^2/\text{s}$ while that of D_n are $1.3 \times 10^{-3} \text{ m}^2/\text{s}$ (Si) and $4.7 \times 10^{-3} \text{ m}^2/\text{s}$ (Ge).

Example 6.7

The mobility of electrons in Si is $0.19 \text{ m}^2/\text{Vs}$, calculate the diffusion coefficient of electrons in Si at 310 K .

Solution.

We know

$$D_n = \frac{\mu_n kT}{e}$$

Here,

$$\mu_n = 0.19 \text{ m}^2/\text{Vs}, T = 310 \text{ K}$$

$$e = 1.6 \times 10^{-19} \text{ C}, k = 1.38 \times 10^{-23}$$

$$\therefore \text{Diffusion coefficient } D_n = \frac{0.19 \times 1.38 \times 10^{-23} \times 310}{1.6 \times 10^{-19}}$$

$$= 5.08 \times 10^{-3} \text{ m}^2/\text{s.}$$

6.14 GENERATION AND RECOMBINATION OF MINORITY CARRIERS

As the temperature of a semiconductor increases, covalent tetrahedral bonds are broken and free electrons are generated. Some of these free electrons may possess sufficient thermal energy to move from valence band to conduction band. A hole is created in the valence band for every free electron moving to valence band. This is referred to as *electron-hole pair generation* (or *production*), as depicted in Fig. 6.13(a).

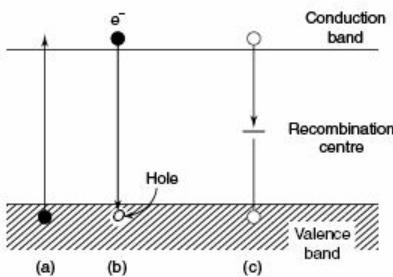


Fig. 6.13 (a) Electron-hole pair generation, (b) direct recombination, (c) recombination through a recombination centre.

In semiconducting materials, electron-hole pairs can be produced by light irradiation, if the frequency ν of the incident is high enough such that $h\nu > E_g$, where E_g is the forbidden energy gap. When the light is removed, the material returns to thermal equilibrium. This occurs by *electron-hole recombination* process. Recombination is said to have taken place, when one of the conducting electrons is bound again on a normal covalent tetrahedral site this is referred to as *band-to-band recombination* since the electron drops from the conduction band into the valence band in the energy band picture, as shown in Fig. 6.13(b). Band-to-band recombination generally results in the emission of photons with energy equal to that of the semiconductor energy gap. In gallium arsenide this type of recombination is common.

In Ge and Si, recombination occurs via a *recombination centre*, which are located at the centre of the energy gap. These levels are introduced by impurity atoms such as Cu and Au.

This type of recombination process may be considered to occur as follows. When a hole jumps from valence band to the recombination centre (also called trap) and an electron from valance band drops to this level, *annihilation* takes place. It may also be considered an electron from the conduction band dropping into a hole in the valence band using recombination level as a stepping stone, as shown in Fig. 6.13(c).

Non-uniformity in periodic structure can occur at an interface between two crystals. Such non-uniformity disrupt the periodic nature of the electric field within the crystal causing regions which become centres of trap or capture of free charge carriers. In the energy band model these disturbances in the electric field create energy levels in energy gap range which has high probability of capturing an electron, a hole or both of them (Fig. 6.14).

These energy levels approaching the conduction band energies are termed *electron traps*, as they have high probability of capturing free electrons. Similarly, those levels near valence band levels are called *hole traps*, having a high probability of capturing a hole. During its life as a carrier, an electron or hole gets trapped several times and being freed to the adjacent band by energy absorption. While trapped the electrons and holes do not contribute to the conduction process. *The average time that an electron or hole remains free after generation is known as its lifetime (τ)*.

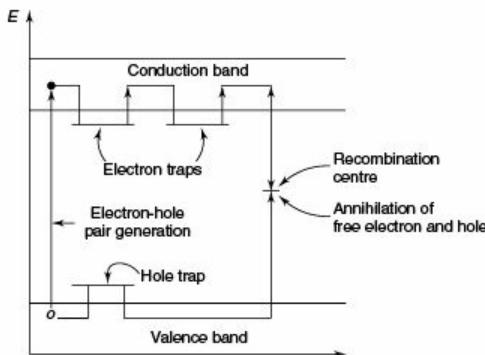


Fig. 6.14 Band model representation of generation and recombination in a semiconductor.

6.15 HALL EFFECT

In 1879, an American Physicist E.H. Hall observed that *when a current-carrying conductor (metal or semiconductor) is placed in a magnetic field perpendicular to the direction of the current, a voltage (called Hall voltage) is developed across the conductor in a direction perpendicular to both the current and magnetic field*. This effect is known as *Hall effect*. This is illustrated in Fig. 6.15. This is the most commonly used to find whether a semiconductor is *n*-type or *p*-type (the nature of charge carrier) and the carrier concentration. Also, by measuring the conductivity σ , the mobility of the charge carrier can be determined.

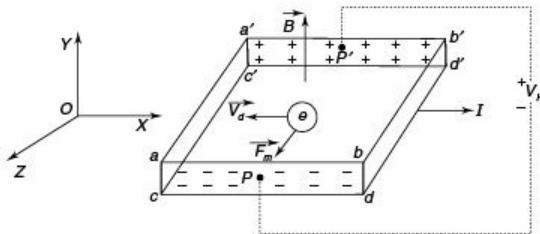


Fig. 6.15 Hall effect.

In Fig. 6.16, current I is in $+x$ direction and the magnetic field B is the $+z$ direction, a force will be exerted in the $-y$ direction on the current carriers. The current I may be due to the holes moving from left to right or free electrons from right to left to right. Hence, irrespective of the sign of the carrier, they will be forced downward towards side 1. If the semiconductor is n -type, then the current is primarily carried by the free electrons and these electrons accumulate on side 1 and this surface becomes negatively charged with respect to side 2. Hence, a potential called the Hall voltage (V_H), appears between surfaces 1 and 2. If the polarity of V_H is positive at terminal on side 2, then the charge carriers must be electrons and the semiconductor is n -type. While if the terminal on side 1 is positive with respect to terminal on side 2, the charge carriers are the holes and the semiconductor must be p -type.

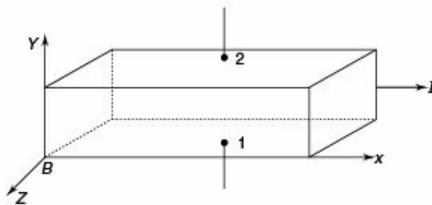


Fig. 6.16

Let an n -type semiconductor slab carrying current I be placed transverse magnetic field B , and v is the velocity of the electrons. The electrons experience a force Bev due to the magnetic field at right angle to both the flow direction and the magnetic field. This deflects the electrons causing a negative charge accumulation on one face of the slab. A potential difference is therefore developed across the surfaces 1 and 2, which in turn creates an electric field E_H . This field exerts a force eE_H on the electrons in the opposite direction. Eventually, an equilibrium occurs when

$$\begin{aligned} eE_H &= Bev \\ \text{or } E_H &= Bv \end{aligned} \quad \dots(6.43)$$

If, J is the current density, then

$$J = nev \quad \dots(6.44)$$

where, n is the concentration of the electrons.

$$\therefore v = \frac{J}{ne}$$

Substituting this in Eq. (6.43), we get

$$E_H = \frac{BJ}{ne} \quad \dots(6.45)$$

A coefficient, called Hall coefficient (R_H), is used to describe the Hall effect. The Hall coefficient R_H is defined as

$$R_H = \frac{1}{ne} \quad \dots(6.46)$$

From Eq. (6.45) it is thus obvious,

$$E_H = R_H B J$$

$$\text{or } R_H = \frac{E_H}{BJ} = \frac{1}{ne} \quad \dots(6.47)$$

Since all the three quantities E_H , J and B are measurable, the Hall coefficient R_H and hence carrier density n can be determined.

Generally, for N-type specimen, the Hall field is developed in negative direction compared to the field developed in a p-type specimen. Therefore, negative sign is used to denote the Hall coefficient, i.e.,

$$R_H = -\frac{E_H}{BJ} = -\frac{1}{ne} \quad \dots(6.48)$$

For p-type specimen when the current is due to the holes,

$$R_H = \frac{E_H}{BJ} = \frac{1}{pe} \quad \dots(6.49)$$

where p is density of holes.

Determination of the Hall Coefficient

Let t be the thickness of the specimen across which the Hall voltage V_H is measured

$$E_H = \frac{V_H}{t}$$

Substituting this for E_H in Eq. (6.49), we get

$$R_H = \frac{V_H}{BJt}$$

or

$$V_H = R_H B J t \quad \dots(6.50)$$

If b is the width of the specimen, so that its cross-sectional area is bt , then the current density is given by

$$J = \frac{I}{bt}$$

$$\therefore V_H = \frac{R_H BI}{b}$$

Hence,

$$R_H = \frac{V_H b}{BI} \quad \dots(6.51)$$

Thus, measuring V_H and knowing B , I and b , the Hall coefficient can be determined. The polarity of V_H will be opposite for n -and p -type semiconductors.

Experimental Arrangement

A slab of the specimen with rectangular surfaces, having a thickness t and width b is placed between two pole pieces of an electromagnet such that the magnetic field B is perpendicular to the direction of the current I flowing in it, as shown in Fig. 6.17. The Hall voltage V_H is then measured by placing two probes at the two side faces of the slab. If the magnetic flux density is B wb/m² and Hall voltage V_H is in volts, then the Hall coefficient R_H can be calculated using Eq. (6.51) in m³/coulomb.

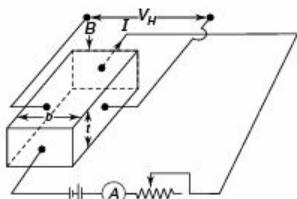


Fig. 6.17 Experimental arrangement for Hall effect experiment.

For n -type specimen, the conductivity is given by

$$\sigma_n = ne \mu_n$$

$$\therefore \mu_n = \frac{\sigma_n}{ne} = -\sigma_n R_H \quad \dots(6.52)$$

For p -type specimen,

$$\mu_n = \sigma_h R_H \quad \dots(6.53)$$

Applications of Hall Effect

(i) Determination of the type of semiconductor

For an n -type semiconductor, the Hall coefficient is negative while for a p -type semiconductor it is positive. Therefore, by measuring the sign of the Hall voltage developed, the type of a given semiconductor, can be determined.

(ii) Determination of carrier concentration, n

By measuring the Hall coefficient R_H as discussed earlier, the carrier concentration can be calculated as

$$n = \frac{1}{e R_H}$$

(iii) Determination of mobility μ

By measuring the conductivity σ and Hall coefficient R_H , the mobility of electron or hole can be calculated.

(iv) *Measurement of magnetic flux density*

Using a semiconducting specimen of known Hall coefficient, the magnetic flux density can be calculated using the relation.

$$R_H = \frac{V_H t}{IB}$$

or

$$B = \frac{V_H t}{IR_H}$$

Magnetic flux density meter is designed on this principle.

(v) *Hall Effect multiplier*

It is an instrument which gives an output proportional to the product of two signals. If current I is made proportional to one of the inputs and if magnetic flux B is linearly related to the second signal, then V_H is proportional to the product of the two inputs.

(vi) *Measurement of power in an electromagnetic wave*

In electromagnetic waves, the magnetic field and electric field are at right angles. Therefore, if a semiconductor in specimen is placed parallel to E it will drive a current I in the semiconductor. As the semiconductor is simultaneously subjected to a transverse magnetic field, Hall voltage develops in the specimen. The Hall voltage will be proportional to the product of E and H , i.e., to the magnitude of the Poynting vector of the electromagnetic wave. Thus, the Hall effect can be used to determine the power flow in an electromagnetic wave.

Example 6.8

A current of 20 A is passed through a Cu slab of 2 cm wide and 0.5 cm thickness. The slab is placed in a magnetic field B of 2.5 T. The magnetic field is perpendicular to the plane of the slab and to the current. Calculate the developed Hall voltage if n for the Cu is 8.48×10^{28} electrons/m³.

Solution.

We know, $R_H = \frac{1}{ne}$ and $V_H = \frac{R_H IB}{t}$

$\therefore V_H = \frac{IB}{net}$

Here, $I = 20 \text{ A}$, $B = 2.5 \text{ T}$

$n = 8.48 \times 10^{28} \text{ electrons/m}^3$, $t = 0.5 \text{ cm} = 0.5 \times 10^{-2} \text{ m}$

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

$$\begin{aligned} \therefore V_H &= \frac{20 \times 2.5}{8.48 \times 10^{28} \times 1.6 \times 10^{-19} \times 0.5 \times 10^{-2}} \text{ volt} \\ &= 7.37 \times 10^{-7} \text{ V} = 0.737 \times 10^{-6} \text{ V} = 0.737 \mu\text{V}. \end{aligned}$$

Example 6.9

A silicon specimen of thickness 1 mm, breadth 10 mm length 100 mm is placed in a transverse magnetic field of 5 Wb/m². If a current of 1A flows along its length, calculate the developed Hall voltage if its Hall coefficient is 3.66×10^{-4} m³/C.

Solution.

$$\text{We know, Hall voltage } V_H = \frac{R_H I B}{t}$$

Here,

$$R_H = 3.66 \times 10^{-4}, \quad I = 1A, \quad B = 5 \text{ Wb/m}^2$$

$$t = 1 \text{ mm} = 10^{-3} \text{ m}$$

$$\therefore V_H = \frac{3.66 \times 10^{-4} \times 1 \times 5}{10^{-3}} = 1.83 \text{ V}$$

Example 6.10

The Hall coefficient of a specimen is 3.66×10^{-4} m³/C.

If its resistivity 8.93×10^{-3} Ωm , calculate its n and μ .

Solution.

We know,

$$R_H = \frac{1}{ne}$$

$$\therefore n = \frac{1}{R_H e} = \frac{1}{3.66 \times 10^{-4} \times 1.6 \times 10^{-19}} = 1.7 \times 10^{22} \text{ m}^{-3}$$

Mobility

$$\mu = \sigma R_H = \frac{R_H}{P}$$

$$\therefore \mu = \frac{3.66 \times 10^{-4}}{8.93 \times 10^{-3}} = 0.04 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$$

Example 6.11

The Hall coefficient of silicon is -7.35×10^{-5} m³ C⁻¹ from room temperature to 400 K. Is this a p-type or n-type semiconductor. If its conductivity is $200 \text{ m}^{-1} \Omega^{-1}$, calculate the mobility and density of the charge carrier.

Solution.

A negative Hall coefficient indicates that the semiconductor is n-type.

As

$$R_H = \frac{1}{ne}$$

$$\therefore \text{Density of charge carrier } n = \frac{1}{R_H e} = \frac{1}{7.35 \times 10^{-5} \times 1.6 \times 10^{-19}} = 8.503 \times 10^{22} \text{ m}^{-3}$$

Conductivity

$$\sigma = ne \mu$$

$$\therefore \text{Mobility} \quad \mu = \frac{\sigma}{ne} = \frac{200}{1.6 \times 10^{-19} \times 8.503 \times 10^{22}} \\ = 1.47 \times 10^{-3} \text{ m}^2 \text{V}^{-1} \text{s}^{-1}$$

Example 6.12

A semiconducting crystal 12 mm long, 5 mm wide and 1 mm thick has a magnetic flux density of 0.5 Wb⁻² transverse to its largest faces. When a current of 20 mA flows lengthwise through the specimen, the voltage measured across its width is 37 µV. Calculate its Hall coefficient.

Solution.

$$\text{Hall coefficient} \quad R_H = \frac{1}{ne} = \frac{E_y}{J_x B_z}$$

$$\text{or} \quad R_H = \frac{V_y}{b J_x B_z} = \frac{V_y}{b} \frac{bt}{I_x B_z} \quad \left(\because E_y = \frac{V_y}{b}, J_x = \frac{I_x}{bt} \right)$$

$$\text{or} \quad R_H = \frac{V_y t}{I_x B_z}$$

Here, $V_y = 37 \times 10^{-6}$ V, $t = 1 \text{ mm} = 10^{-3} \text{ m}$, $B_z = 0.5 \text{ Wb m}^{-2}$

$$I_x = 20 \times 10^{-3} \text{ A}$$

$$\therefore R_H = \frac{37 \times 10^{-6} \times 10^{-3}}{20 \times 10^{-3} \times 0.5} = 3.7 \times 10^{-6} \text{ C}^{-1} \text{m}^3$$

6.16 FORMATION OF P-N JUNCTION DIODE

A *p*-type or *n*-type silicon (or germanium) can be grown by adding suitable and controlled amount of impurity in the melt of pure silicon (or germanium). The crystal grown are cut into thin slices called wafer. Semiconductor devices such as diode, transistor, field effect transistor (FET), etc. are grown on these wafers.

If on a wafer of *n*-type silicon, a thin film of trivalent material, say aluminium, is placed and heated to a high temperature ($\approx 575^\circ \text{C}$), aluminium diffuses into silicon. Consequently, a *p*-type semiconductor is formed on an *n*-type semiconductor. The boundary (or surface of contact or region of transition between *p*-type semiconducting materials) is called the *p-n* junction. A *p-n* junction can also be grown by diffusing a pentavalent material, say phosphorus, into a *p*-type semiconductor.

The wafer on which *p-n* junction are grown, is cut into small pieces. The simplest semiconductor device is *p-n* junction diode. A *p-n* junction diode consists of *p* and *n* semiconducting regions on a single wafer encapsulated in a casing with electrical contact coming out from *p* and *n* regions.

There are numerous ways of forming *p-n* junction and it is also possible to form more than one junction (e.g., *pnp* and *npn*) on the same wafer. We shall not go here into the details of these processes.

Figure 6.18(a) shows a *p-n* junction diode, *p*-side is also known as anode and *n*-side cathode. The symbolic representation of a *p-n* junction diode is shown in Fig. 6.18(b). The arrow indicates the direction of conventional current flow.

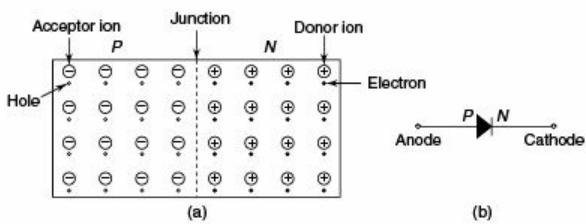


Fig. 6.18 (a) Formation of a *p-n* junction. (b) Symbol of a *p-n* junction diode.

6.17 DEPLETION LAYER AND BARRIER POTENTIAL IN A P-N JUNCTION DIODE

We discuss below the physics involved in a *p-n* junction diode.

In a *p-n* junction diode, the *p*-region has high concentration of holes and the *n*-region that of electrons. The moment *p-n* junction is formed, holes from the *p*-region start diffusing across the junction into the *n*-region while electrons from the *n*-region diffuse across the junction into the *n*-region. When an electron diffuses into *p*-region, it (the electron) falls into a vacancy, in other words, it recombines with a hole. This process is called *recombination*. This recombination results in the production of an ionized acceptor atom, as it has accepted an electron. The ionized acceptor atom becomes negatively charged and is immobile. In the *p*-region, in the immediate vicinity of the junction, immobile excess negative charges are formed.

Similarly, holes from the *p*-region recombine with electrons in *n*-side on crossing the junction. This results in the formation of immobile excess positive charges in the vicinity of the junction on the *n*-side.

These layers of (immobile) charge create an *electric field*, which exerts a force on the electrons and holes and thus opposes their further diffusion across the junction (Fig. 6.19). The electric field becomes strong enough, as diffusion proceeds, to stop it. Due to this field, there is a potential at the junction (with *n* side at higher potential), known as *barrier potential* (V_B). This is so called because this potential acts as a barrier for charge motion in the diffusion process. The magnitude of the barrier potential is 0.3 V for germanium junction diode and about 0.7 V for silicon junction diode. Besides the nature of the crystal, the amount of doping and temperature of the crystal are the other factors which determine the magnitude of V_B .

Because of inter-diffusion, the region in the vicinity of the junction has very low concentration of charge carriers. This region is called *depletion region*. The thickness of this region is about $1 \mu\text{m}$ (10^{-6} m). The barrier potential is the potential difference across the depletion with *n*-side of the depletion region at higher potential (Fig. 6.19). A very large electric field is set up across the depletion region. The approximate strength of the electric field is

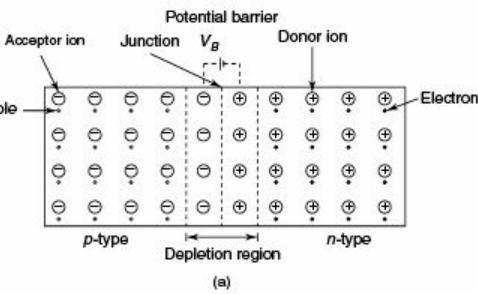


Fig. 6.19

$$E = \frac{\text{Barrier potential}}{\text{Thickness of depletion region}} = \frac{0.7 \text{ V}}{10^{-6} \text{ m}} = 7 \times 10^3 \text{ Vm}^{-1}$$

6.18 BIASING OF P-N JUNCTION DIODE

Applying a suitable external potential difference to a *p-n* junction diode is called the *biasing*. There are two ways in which a *p-n* junction diode can be biased. Both of these are discussed below.

- (i) **Forward Biasing:** When a battery is connected to a *p-n* junction diode in such a way that positive of the battery is connected to *p*-side and negative to *n*-side, the junction diode is said to be forward biased.

If the diode is forward biased, the externally applied voltage creates an electric field across the *p-n* junction that is opposite to the field due to potential barrier (V_B) (Fig. 6.20). Therefore, the potential across the junction decreases, i.e., potential barrier is lowered and the diffusion of electrons and holes across the junction thereby increases, resulting in a current in the circuit. In case the forward bias (V) is greater than the barrier potential (V_B), charge carriers, i.e., electrons and holes diffuse easily resulting in a sharp increase in the current through the junction diode.

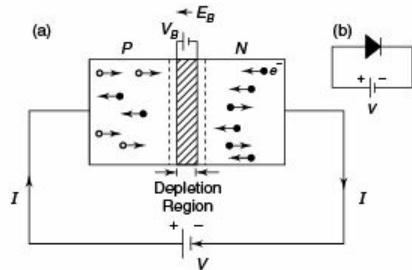


Fig. 6.20 A forward biased *p-n* junction diode. The dotted lines indicate original (unbiased) depletion layer. V_B —barrier potential, V —applied potential.

Due to motion of majority carrier, that is, holes from *p*-region and electrons from *n*-region, towards the junction in forward bias condition the thickness of the depletion region decreases or depletion region shrinks. A reduced depletion region offers a low resistance to the flow of current. The current flowing in the circuit when the junction diode is forward biased is also known as forward current.

- (ii) **Reverse Biasing:** If a battery is connected to *p-n* junction diode with positive terminal of its (battery) joined to the *n*-side of the junction diode and negative terminal joined to the *p*-side of the junction diode, then the *p-n* junction diode is said to be reverse biased. In case of reverse biasing, the externally applied potential and barrier potential are in same direction. Due to this thickness of depletion region increases or resultant barrier potential increases inhibiting the diffusion of holes and electrons across the junction (Fig. 6.21). The resistance offered by the depletion region becomes very high and consequently current drops to a negligibly small value. However, a very small amount of current flows due to

the motion of minority carriers, i.e., holes in *n*-region and electrons in *p*-region. This current is called *reverse current*.

Potential barrier in the junction region opposes the flow of majority carriers but accelerates minority carriers.

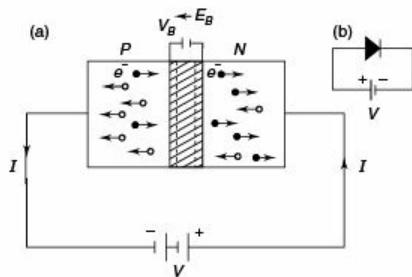


Fig. 6.21 A reverse-biased *p-n* junction diode. The dotted lines indicate original depletion region.

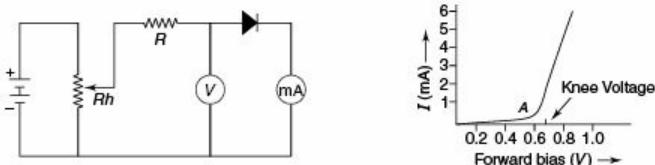
6.19 CHARACTERISTICS OF A P-N JUNCTION DIODE

The relation between applied voltage (V) and the resulting current (I), expressed graphically, is known as characteristic curve of a device. In a *p-n* junction diode the characteristic curve in forward bias condition differs drastically from that in reverse bias condition. We discuss the two characteristics below.

6.19.1 Forward Bias Characteristics

Figure 6.22(a) shows the circuit diagram for obtaining the characteristic curve of a forward biased *p-n* junction diode. The battery is connected to the diode through a rheostat (or potentiometer) so that the voltage applied to the diode can be changed. The voltmeter across the diode measures the voltage across the diode and the milliammeter measures the current in the diode. The voltage across the diode is increased in small steps, say 0.1 V, and the corresponding current is noted. A typical forward characteristic curve for a silicon diode ($V_B \approx 0.7$ V) is shown in Fig. 6.22(b).

As shown in Fig. 6.22(b), the current increases very slowly across the diode till a certain voltage, a characteristic of the diode, is reached. After this characteristic voltage, the diode current increases rapidly, even for very small increase in the diode bias voltage. This voltage is called the *threshold voltage* or *cut-in voltage* or *knee voltage*. The value of this voltage is about 0.3 V for a germanium diode and 0.7 V for a silicon diode.



(a) Circuit diagram for characteristic curve of a forward biased *p-n* junction diode.

(b) Characteristic curve of a forward biased *p-n* junction diode.

Fig. 6.22

To understand the characteristic curve, it is important to note that till the applied potential is less than the barrier potential, there is very small current because the barrier potential exists. As the applied potential is increased the barrier potential decreases and initially current increases slowly. When the applied potential becomes equal or exceeds the barrier potential, the thickness of depletion region becomes negligible. Electrons from *n*-region and holes from *p*-region, move across the junction without encountering any appreciable resistance. As a consequence, a relatively large current flow through the diode.

6.19.2 Reverse Bias Characteristics

Figure 6.23(a) shows the circuit diagram for obtaining the characteristic curve of a reverse biased *p-n* junction diode. Here a microammeter (μA) is used. A typical reverse characteristic curve is shown in Fig. 6.23(b).

When the diode is reverse biased, the reverse bias voltage produces an extremely small current, about a few microamperes (μA) which remains nearly constant. This current is also sometimes referred to as *leakage current*. However, when the applied reverse voltage reaches a characteristic voltage called the *breakdown voltage* the current suddenly increases to a large value. The breakdown voltage is also called the *peak-inverse voltage* of the diode. A normal *p-n* junction diode (unlike Zener diode, discussed later) is not operated in the reverse bias beyond the breakdown voltage.

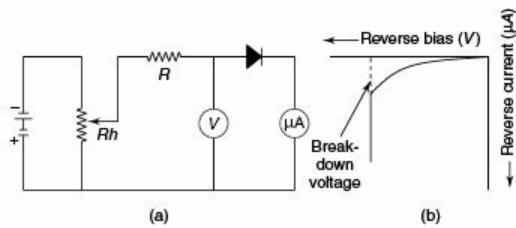


Fig. 6.23 (a) Circuit diagram for characteristics of a reverse biased *p-n* junction diode
(b) Characteristic curve of a reverse biased *p-n* junction diode.

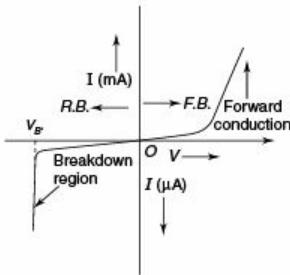


Fig. 6.24 A typical complete diode characteristic in the forward and reverse region (not to scale).

6.19.3 Resistance of Junction Diode

The current-voltage curve of a *p-n* junction diode is not a straight line, i.e., it does not obey Ohm's law and thus it is a *non-ohmic device*. Further, the resistance offered by the junction diode depends upon the applied voltage due to the characteristic curve not being a straight line.

The d.c. resistance ($R_{d.c.}$) of a *p-n* junction diode is the ratio of the applied voltage (V) and the resulting current (I), that is,

$$R_{d.c.} = \frac{V}{I}$$

The d.c. resistance is not a very useful parameter. The dynamic or a.c. resistance ($R_{a.c.}$) is the most commonly used parameter. The dynamic resistance of a junction diode is defined as the ratio of small change in voltage (ΔV) to the small change in current produced (ΔI). That is,

$$R_{a.c.} = \frac{\Delta V}{\Delta I}$$

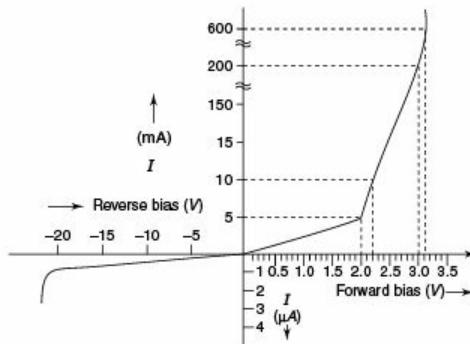


Fig. 6.25

The region of the characteristic curve where the dynamic resistance is almost independent of the applied voltage is called the linear region of junction diode.

Example 6.13

The $I-V$ characteristic of a *p-n* junction diode is shown in Fig. 6.25. Find the approximate dynamic resistance of the *p-n* junction when (i) a forward bias of 2 volts is applied, (ii) a forward bias of 3 volts is applied.

Solution.

- The current at 2 V is 5 mA and at 2.2 V it is 10 mA. The dynamic (or a.c.) resistance in this region is

$$R_{\text{a.c.}} = \frac{\Delta V}{\Delta I} = \frac{2.2 - 2.0 \text{ volt}}{10 - 5 \text{ mA}}$$

$$= \frac{0.2}{5} \times 10^3 \Omega \\ = 40 \Omega$$

- (ii) The current at 3 volts is 200 mA and at 3.1 volts is 600 mA. The dynamic resistance in this region is

$$R_{\text{a.c.}} = \frac{\Delta V}{\Delta I} = \frac{3.1 - 3.0 \text{ V}}{600 - 200 \text{ mA}} = \frac{0.1}{400} \times 10^3 \Omega = 0.25 \Omega$$

Example 6.14

In the above problem, determine the resistance of the diode when the current is (i) 5 mA, (ii) 200 mA and (iii) $V = -10 \text{ V}$.

Solution.

(i) When $I = 5 \text{ mA}, V = 2 \text{ volt}$

$$\therefore \text{Resistance } R_{\text{d.c.}} = \frac{2 \text{ V}}{5 \times 10^{-3} \text{ A}} = \frac{2}{5} \times 10^3 \Omega = 400 \Omega$$

(ii) When $I = 200 \text{ mA}, V = 3 \text{ volt}$

$$\therefore \text{Resistance } R_{\text{d.c.}} = \frac{3 \text{ V}}{200 \text{ mA}} = \frac{3}{2} \times 10 \Omega = 15 \Omega$$

(iii) When $V = -10 \text{ volt}, I = -1 \mu\text{A} = -10^{-6} \text{ A}$

$$\therefore \text{Resistance } R_{\text{d.c.}} = \frac{-10 \text{ V}}{-10^{-6} \text{ A}} = 10 \times 10^6 \Omega = 10 \text{ M}\Omega$$

Example 6.15

In the following figures (Fig. 6.26) which of the diodes are forward biased and which are reverse and why?

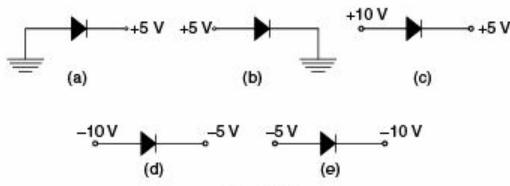


Fig. 6.26

Solution

- The *p*-side of the junction diode is grounded, i.e., it is at zero potential while the cathode is at a potential of + 5 V. Since the *n*-side (cathode) is at higher potential than *p*-side (anode), the junction diode is *reverse biased*.
- Since the *p*-side is at higher potential (+ 5 V) than the *n*-side (0 V) the junction diode is *forward biased*.
- Since the *p*-side is at higher potential (+ 10 V) than the *n*-side (+ 5 V) the junction diode is *forward biased*.
- Since the *n*-side is at higher potential (- 5 V) than the *p*-side (- 10 V) the junction diode is *reverse biased*.
- Since the *n*-side is at lower potential (- 8 V) than the *p*-side (- 5 V) the junction diode is *forward biased*.

6.20 ZENER DIODE

In Fig. 6.27, the I - V characteristic of a *p*-*n* junction diode including the diode which are meant to be operated in reverse bias condition, such as a Zener diode, is shown. As seen in the figure, if the reverse bias voltage across a *p*-*n* junction diode is increased, at a particular voltage called *breakdown voltage*, the (reverse) current suddenly increases to a large value. *The Zener diode is a specially designed junction diode to operate in the breakdown region without getting damaged.* The symbol of a Zener diode is shown in Fig. 6.28(a).

Such a diode can be used as a *constant-voltage device*, as shown in Fig. 6.28(b).

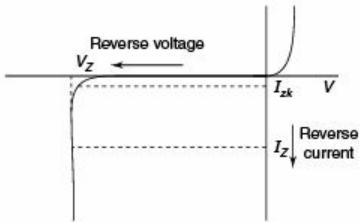


Fig. 6.27 V - I characteristic of a Zener diode.

The voltage at which the breakdown occurs is known as the zener (or breakdown) voltage (V_z) for the Zener diode. The breakdown voltage depends upon the extent of doping. A Zener diode is comparatively heavily doped so that the breakdown occurs at a lower voltage. Zener diodes with breakdown voltage between 5 V to 200 V are commercially available.

When the reverse voltage across the Zener diode is equal or more than its breakdown V_z , the current increases sharply. After the breakdown, a small change in voltage produces a very large change in current or in other words corresponding to a large change in voltage takes place. This property is utilized in using Zener diode as *voltage regulator*.

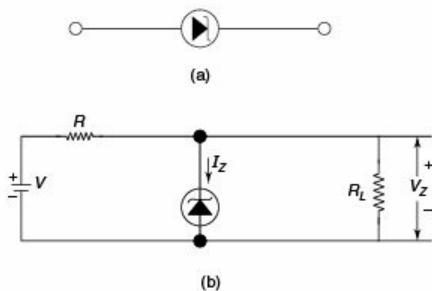


Fig. 6.28 (a) Symbol of a Zener diode (b) Zener diode as a voltage regulator.

6.20.1 Zener Diode as Voltage Regulator

A Zener diode can be used as a constant-voltage device or voltage regulator using a circuit given in Fig. 6.28(b). When the input voltage V and resistor R is such that voltage across the Zener diode is equal to or greater than its breakdown voltage V_Z , the Zener diode operates in the breakdown region. The voltage across the load R_L is the same as across the diode, V_Z and the diode current is I_Z . The diode now can regulate (or keep constant) the load voltage even if there is a variation in the load current and supply (input) voltage. If there is a change in load voltage or supply voltage, the diode current adjusts itself (increase or decrease) so as to maintain a nearly constant load voltage. The diode continues to act as a voltage stabilizer until the circuit changes require the diode current to fall to I_{ZK} , near the knee of the diode $I-V$ curve. The upper limit of the diode current is determined by the power-dissipation capability of the diode.

6.20.2 Difference Between Zener Breakdown and Avalanche Breakdown

In normal $p-n$ junction diodes, the breakdown takes place by avalanche breakdown. The avalanche breakdown takes place as discussed below. When a thermally generated electron (part of reverse current) falls down the junction barrier and acquires energy from the applied potential, it may collide with a crystal ion breaking the covalent bond and thus creating a new electron-hole pair. These carriers in turn acquire sufficient energy to collide with another crystal ion generating another electron-hole pair. This cumulative process leads to avalanche multiplication. This results in a large reverse current in the diode and the diode is to be in the region of *avalanche breakdown*.

In Zener diode, the breakdown takes place by *Zener breakdown* which is different from avalanche breakdown. A Zener diode is comparatively heavily doped. Because of the existence of a comparative strong electric field at the junction, a sufficiently strong force may be exerted on a bound electron, forcing one of its covalent bond, generating a new electron-hole pair and so increasing the reverse current when this process continues, the reverse current increases sharply. This is known as *zener breakdown*. Zener breakdown occurs at an electric field of approximately $2 \times 10^7 \text{ Vm}^{-1}$. This value of the field is reached for voltage below 6 V for heavily doped diodes. For lightly doped diodes the breakdown voltage is higher and the avalanche multiplication is a predominant effect.

Example 6.16

In Fig. 6.29, what is the voltage required to maintain 15 V across the load resistance R_L of $2 \text{ k}\Omega$, assuming that the series resistance R is 200Ω and the Zener diode requires a minimum current 10 mA to work satisfactorily? What is the Zener rating required?

Solution.

The current distribution between Zener and load R_L is shown in Fig. 6.29

Given: Load resistance $R_L = 2 \text{ k}\Omega = 2 \times 10^3 \Omega$

Current required for a voltage of 15 V across R_L ,

$$I_L = \frac{V_L}{R_L} = \frac{15 \text{ V}}{2 \times 10^3} = 7.5 \times 10^{-3} \text{ A}$$

$$= 7.5 \text{ mA}$$

Current required in Zener diode = 10 mA

$$\therefore \text{current through } R, I_R = I_Z + I_L = 10 + 7.5 \\ = 17.5 \text{ mA}$$

Series resistance $R = 200 \Omega$

Voltage across $R, V_R = R \cdot I_R = 200 \times 17.5 \times 10^{-3} = 3.5 \text{ V}$

$$\therefore \text{Input voltage needed, } V = V_R + V_L = 3.5 + 15 = 18.5 \text{ V}$$

The Zener diode rating should be current 17.5 mA, and breakdown voltage = 15 V

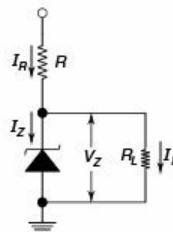


Fig. 6.29

Example 6.17

In the circuit shown in Fig. 6.30, calculate (i) output voltage, voltage across series resistance and (ii) current through Zener diode.

Solution.

In the absence of the Zener diode, the open circuit voltage is

$$V = \frac{R_L V_i}{R + R_L} = \frac{10 \times 10^3 \times 120}{5 \times 10^3 + 10 \times 10^3} \\ = 80 \text{ V}$$

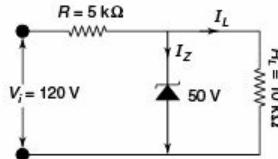


Fig. 6.30

Since the breakdown voltage of the Zener diode is 50 volts, the diode will undergo the breakdown.

(i) Output voltage $V_L = 50 \text{ V}$

(ii) Voltage drop across $R, V_R = 120 - 50 = 70 \text{ V}$

(iii) Current through the load, $I_L = \frac{50 \text{ V}}{10 \text{ k}\Omega} = 5 \text{ mA}$

$$\text{Current through } R, I_R = \frac{70 \text{ V}}{5 \text{ k}\Omega} = 14 \text{ mA}$$

$$\therefore \text{Current through Zener diode, } I_Z = I_R - I_L = 14 - 5 \text{ mA} \\ = 9 \text{ mA}$$

Example 6.18

In the circuit diagram shown in Fig. 6.31, a 6.5 V Zener diode is used. The input voltage of 12 V is applied and the minimum Zener current is 10 mA. If the load current is to vary from 10 to 100 mA. Find the value of the series resistance R to maintain a voltage of 6.5 V across the load.

Solution.

$$V_i = 12 \text{ V}, V_Z = 6.5 \text{ V}$$

$$R = \frac{V_i - V_o}{I_Z + I_L}$$

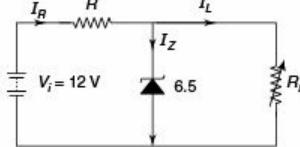


Fig. 6.31

The voltage across R is to remain constant at $12 - 6.5 = 5.5 \text{ V}$. The minimum Zener current will be there when the load current is maximum.

$$R = \frac{V_i - V_o}{I_{z \min} + I_{L \max}} = \frac{12 - 6.5 \text{ V}}{(10 + 100) \text{ mA}} = \frac{5.5 \text{ V}}{110 \text{ mA}} = 50 \Omega$$

If $R_S = 50 \Omega$, The output will remain constant.

6.21 TUNNEL DIODE

You might have come across in quantum mechanics that there is a finite probability that a particle can ‘tunnel’ through a barrier if the barrier is thin enough. This is in contradiction to classical idea, where a particle must have energy greater than or equal to the potential energy barrier if it is to move across the barrier. The quantum mechanical process is referred to as tunnelling.

A diode based on this quantum mechanical process of tunnelling was developed by a Japanese physicist Leo Esaki in 1958. He shared Nobel Prize in Physics 1973 with Ivar Giaever and Brian David Josephson for the discovery of electron tunnelling.

Construction. The width of the depletion layer (or junction barrier) varies inversely with the square root of impurity concentration. In a normal $p-n$ junction diode, with an impurity concentration of about 1 part in 10^8 , the width of depletion layer is of the order of one micron (10^{-6} m). The corresponding potential barrier restricts the flow of respective majority carriers (holes in p -region and electrons in n -region) across the junction. However, if the concentration of impurity atoms is increased considerably, say, to 1 part in 10^3 , the width of depletion region decreases significantly ($\approx 10 \text{ \AA}$) and the characteristics of the so obtained $p-n$ junction diode changes drastically. A tunnel diode is based on this concept.

Thus, a tunnel diode is a $p-n$ junction diode in which the two regions are heavily doped (≈ 1 part in 10^3) resulting in an extremely thin depletion layer ($\approx 100 \text{ \AA}$ or 10^{-6} cm). Due to the extremely thin depletion layer, the electrons are able to tunnel through the potential barrier at relatively low forward bias voltage, resulting in an interesting V-I characteristic. These diodes are generally fabricated using germanium, gallium-arsenide (Ga As) and gallium-antimonide (Ga Sb). It is difficult to make a silicon tunnel diode with good characteristics.

V-I Characteristics: The volt-ampere characteristic of a tunnel diode is shown in Fig. 6.32. It is an excellent conductor in reverse bias condition. Also, for a small forward bias (upto 50 mV for Ge), it has small resistance ($\approx 5 \Omega$) and therefore the current increases sharply with small forward bias. The current rises quickly to its peak value, called *peak current*, I_p for the forward voltage V_p (point-A). If the voltage is increased beyond V_p , the current decreases to a minimum value I_v (point B). The minimum current I_v is called *valley current* corresponding to valley voltage V_v . Beyond this point the current increases again with increasing forward bias as in a normal p-n junction diode.

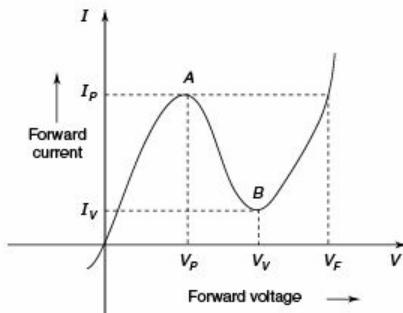


Fig. 6.32 V-I characteristic of a tunnel diode.

The tunnel diode exhibits a *negative-resistance characteristic* between the points A and B. In this region, the current decreases on increasing the voltage. Thus, a tunnel diode acts as negative resistor ($-R_N$) between the points A and B. In this region, the current keeps on decreasing with increasing voltage, meaning that the diode negative resistance increases upto the valley point B. It can be noted in Fig. 6.32 that for current between I_p and I_v , the curve is triple-valued, i.e., a current can be obtained for three values of applied voltage. This feature makes tunnel diode useful in pulse and digital circuits.

The symbol and equivalent circuit of a tunnel diode is shown in Fig. 6.33(a) and (b) respectively. In the equivalent circuit, the series resistance R_s is the ohmic resistance. The series inductance L_s depends upon the length of the connecting wires and the geometry of the dipole package. The bias applied determines the junction capacitance C and is normally measured at the valley point. For a typical tunnel diode with $I_p = 10 \text{ mA}$, the values of the characteristic parameters are: $R_N = -30 \Omega$, $R_s = 1 \Omega$, $L_s = 5 \text{ nH}$ and $C = 20 \text{ pF}$.

The interesting property of a tunnel diode is that in the negative resistance region, it produces power unlike the absorption of the power by a usual (positive) resistor. A tunnel diode can act as high-speed switch. As tunnelling takes place at the speed comparable to the speed of light, the transient response is limited by total shunt resistance (sum of junction capacitance and stray wiring capacitances) switching times as low as 50 ps have been obtained although generally it lies in nanosecond range. Tunnel diodes are used in high frequency (microwave) oscillator. The negative resistance produces oscillations in the tank circuit by offsetting losses in the L and C components.

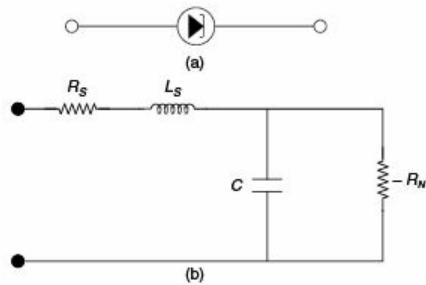


Fig. 6.33 (a) Symbol of a tunnel diode, (b) Equivalent circuit of a tunnel diode.

Tunnelling process: In quantum mechanics you might have come across the details of tunnelling process. In a tunnel diode this process is depicted in Fig. 6.34. With no forward biasing, the energy levels of the conduction electron in n -region not aligned (at different levels) with the energy levels of the holes in the p -region. When a small forward bias is applied, the electron levels tend to get aligned with the hole levels on the other side of the junction, making it possible for the electrons to tunnel through the junction. As the voltage is increased to V_p , all the conduction band electrons in the n -region are able to cross-over to the valence band in the p -region as the two bands across the junctions are exactly aligned. Therefore, a maximum current (peak current I_p) flows in the circuit. As the forward bias is increased beyond V_p , the two bands start getting out of alignment and the current starts decreasing since the tunnelling decreases. The current reaches a minimum value (valley current I_v) when the two levels get totally out of alignment. For forward bias beyond V_p , the current increases as in a normal $p-n$ junction diode.

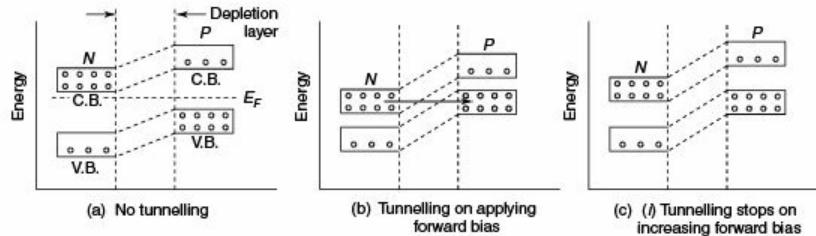


Fig. 6.34

Tunnel Diode Parameters:

- (i) Negative Resistance ($-R_N$). The negative resistance is given by

$$R_N = - \frac{dV}{dI}$$

It is the reciprocal of the slope of the curve in the negative resistance region.

For a typical tunnel diode $-R_N \approx 30 \Omega$

- (ii) I_P/I_V Ratio: This is an important parameter. For a Ge diode, $\frac{I_P}{I_V} = 8$ and that for Ga As diode, $\frac{I_P}{I_V} = 15$. For a Si diode, it is 3.5.

The peak point (V_P, I_P) is not a sensitive function of temperature. However, the valley point (V_V, I_V) is quite temperature dependent.

Applications of Tunnel Diode

Tunnel diode is used for the following applications.

- (i) It is used as a high-speed switch. Switching time of nanosecond (or even picosecond) can be achieved. This is possible because tunnelling takes place at a speed comparable to the speed of light.
- (ii) It is used in a high frequency (microwave frequencies) oscillator.
- (iii) Due to its multivalued values of I and V , it is used in digital (on-off) device.

Advantages of Tunnel Diode

Its main advantages are:

- (i) low noise,
- (ii) low power operation,
- (iii) high-speed,
- (iv) environmental immunity and
- (v) low cost.

Disadvantages of Tunnel Diode

Its few disadvantages are:

- (i) Low output voltage swing
- (ii) It being a two-terminal device due to which there is no isolation between input and output leading to difficulties in circuit design.

SHORT ANSWER TYPE QUESTIONS

1. Distinguish between energy level and energy band.
2. On the basis of energy band diagrams, differentiate between a conductor, semiconductor and insulator.
3. Distinguish between intrinsic and extrinsic semiconductors.
4. A semiconductor is an insulator at 0K. Defend the statement.
5. Distinguish between *n*-type and *p*-type semiconductors.
6. How the conductivity of a semiconductor changes with the increase in temperature? Explain.

7. What do you mean by 'mobility' of a charge carrier in a semiconductor? Obtain its unit.
8. How the conductivity of a semiconductor is dependent on the mobility of charge carriers? Explain.
9. Differentiate between drift and diffusion currents. Name a respective parameter which is related to these currents and define them.
10. What do you mean by (i) pair production, (ii) recombination, (iii) recombination centre, (iv) annihilation and (v) life time of a carrier?
11. How are the electron and hole traps formed? Explain.
12. What is Hall effect? Explain its cause and its significance.
13. Define: (i) Hall voltage and (ii) Hall coefficient.
14. Discuss the applications of Hall effect.
15. Explain: (i) depletion layer and (ii) barrier potential.
16. What is a Zener diode? How does it differ from an ordinary diode?
17. Differentiate between Zener and avalanche breakdowns.
18. What is a tunnel diode? Draw its I-V curve and explain negative resistance.
19. Name and define important parameters of a tunnel diode.
20. What are the applications of a tunnel diode? Explain.

LONG ANSWER TYPE QUESTIONS

1. Explain how energy bands are formed in solids. Differentiate between conduction band and valence band.
2. Discuss the process of conduction in intrinsic semiconductors.
3. Differentiate between *n*-type and *p*-type semiconductors on the basis of energy band diagram. Explain the process of conduction in both types of materials.
4. Using the concept of electron and hole currents, derive expression for the conductivity of the semiconductor.
5. Obtain expression for concentration of electrons and holes in a semiconductor.
6. What do you mean by intrinsic carrier concentration? Obtain an expression for the same. Show that the Fermi level in an intrinsic semiconductor is situated in the middle of the valence and conduction bands.
7. What are the currents in a semiconductor? Obtain an expression for the total current in a semiconductor.
8. Obtain Einstein's equation and explain its significance.
9. Discuss Hall effect and obtain an expression for the Hall coefficient in terms of current *I* and magnetic fields *B*.
10. Discuss the construction and working of a Zener diode. Explain how it can work as a voltage regulator.
11. Discuss the construction and working of a tunnel diode. Mention its applications and comparative advantages and disadvantages.

NUMERICAL EXERCISES

1. For silicon with E_g of 1.12 eV, determine the Fermi level at 300 K if $m_e^* = 0.12 m_o$, and $m_h^* = 0.28 m_o$

$$\left(\text{Hint: Use } E_F = \frac{E_g}{2} + \frac{3kT}{4} \left(\ln e_n \frac{m_h^*}{m_e^*} \right) \text{ to obtain 0.576 eV} \right)$$

2. For Ge at 300 K: $n_i = 2.4 \times 10^{19} \text{ m}^{-3}$, $\mu_e = 0.39 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$, $\mu_p = 0.19 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. What is its resistivity?
(0.449 $\Omega \text{ m}$)
3. For a Ge rod of 1 m length, 1 mm width and 1 mm thickness at 300 K: $n_i = 2.5 \times 10^{19} \text{ m}^{-3}$, $\mu_e = 0.39 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ and $\mu_p = 0.19 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. Find its resistance.
(4.31 k Ω)
4. For a Si with E_g of 1.1 eV, calculate n_i and σ_p . Given: $\mu_e = 0.48 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ and $\mu_h = 0.013 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$.
[Ans. $1.63 \times 10^{16} \text{ m}^{-3}$, $1.287 \times 10^{-3} \Omega^{-1} \text{ m}^{-1}$]
5. In a Hall experiment, a current of 25 A is passed through a long foil of silver which is 0.1 mm thick and 3 cm wide. If the magnetic field of flux density 0.14 Wb/m² is applied perpendicular to the foil, calculate the Hall voltage developed and estimate the mobility of electrons in silver. Given: σ (silver) = $6.8 \times 10^7 \Omega^{-1} \text{ m}^{-1}$ and R_H (silver) = $-8.4 \times 10^{-11} \text{ m}^{-3} \text{ C}^{-1}$.
6. If the carrier mobility of a sample is $0.2 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$ at 0°C, what is the diffusion constant at the same temperature.
(0.0047 $\text{m}^2 \text{ s}^{-1}$)
7. The resistivity of a doped silicon sample is $8.9 \times 10^{-3} \Omega \text{ m}$. The Hall coefficient was measured to be $3.6 \times 10^{-4} \text{ m}^{-3} \text{ C}^{-1}$. Assuming single carrier conduction, find the mobility and density of charge carriers.
(0.0344 $\text{m}^2 \text{ V}^{-1} \text{ s}^{-1}$, $2.04 \times 10^{22} \text{ m}^{-3}$)

7

Dielectric, Ferroelectric and Piezoelectric Properties of Materials

7.1 DIELECTRIC MATERIALS: INTRODUCTION

The materials devoid of free charges are called dielectrics or insulators. An ideal dielectric is one which is completely devoid of free charges. On the basis of energy band diagram, these are the materials with forbidden gap energy of ≥ 3 eV. The large energy gap makes it difficult for the electrons in the valence band to get excited to the conduction band. Consequently, these materials are poor conductors of electricity. The resistivity of most dielectrics lie in the range of $10^{12} \Omega \text{ m}$ or higher, e.g., mica – $10^{11} \Omega \text{ m}$, porcelain – $10^{12} \Omega \text{ m}$, pyrex glass – $10^{14} \Omega \text{ m}$.

Dielectric materials are widely used for physical separation between the conductors. However, the ability of these materials to get polarized under the influence of an applied electric field makes them useful for a wide variety of devices, as discussed later in this chapter. When an external electric field is applied on a dielectric, the atoms and molecules of the dielectric get influenced by the field. The positively charged particles are pushed in the direction of the field while the electrons and other negatively charged particles are pushed in the opposite direction, from their equilibrium positions. Thus, the electric dipoles are formed and they produce their own electric field. The process of producing electric dipoles out of neutral atoms and molecules is referred to as *polarization*.

Dielectric materials are classified as *non-polar dielectrics* and *polar dielectrics*. Generally, atoms and molecules have an equal number of electrons and protons. If the molecule has symmetric structure, the centres of positive and negative charge coincide and the molecule has no net charge and a zero dipole moment. Such symmetric molecules (e.g., H₂, N₂, O₂, CO₂, CH₄, etc.) are known as non-polar molecules and the dielectrics comprising such molecules are called *non-polar dielectrics*. On the other hand, the molecules whose centres of positive and negative charges do not coincide have a net dipole moment. Such molecules are called polar-molecules (e.g., HCl, NH₃, H₂O, N₂O, CO, CH₃OH, etc.) and the dielectrics comprising such molecules are called *polar dielectrics*. In the absence of an applied electric field, the electric dipole movement of the polar molecules are randomly oriented and cancel each other. So even though each molecule has a dipole moment, the average dipole moment per unit volume is zero.

7.2 REVIEW OF BASICS

Let us recall some of the basic terms frequently used in dielectrics.

On the application of an external electric field the positive and negative charges move in opposite directions, creating electric dipoles. An electric dipole consists of two equal and opposite ($-q$ and $+q$) charges separated by a small distance, as shown in Fig. 7.1. The *dipole moment* p is defined as.

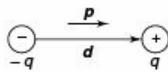


Fig. 7.1 An electric dipole.

$$p = q d \quad \dots(7.1)$$

where, d is the separation between the two charges, from negative to positive charge (a vector quantity). The dipole moment is a vector quantity and its direction is from negative to positive charge. Debye is the unit of electric dipole moment (1 Debye = 3.33×10^{-30} cm). An electric dipole produces an electric field around it, which at a distance r from the dipole moment p is given by

$$E(r) = \frac{3(p \cdot r)r - r^2 p}{4\pi \epsilon_0 r^5} \quad \dots(7.2)$$

In the derivation of the above expression, it is assumed $r \gg d$ and so the expression is strictly valid at a point far off from the dipole. In atoms and molecules, however, this condition is well satisfied as d , being of the order of the atomic diameter, is indeed very small.

When a dipole is placed in an electric field E , a torque acts on the dipole given by

$$\tau = p \times E \quad \dots(7.3)$$

or $\tau = pE \sin \theta \quad \dots(7.4)$

where θ is the angle between the directions of the electric field and the dipole moment (Fig. 7.2). The torque tends to align the dipole in the direction of the field. The interaction of the dipole with an electric field is also expressed in terms of potential energy of the dipole. The potential energy of a dipole is given by

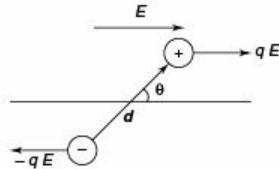


Fig. 7.2 Torque exerted by field E on an electric dipole.

$$U = -p \cdot E \quad \dots(7.5)$$

or $U = -pE \cos \theta \quad \dots(7.5)$

Thus, the potential energy varies from $-pE$ for $\theta = 0^\circ$ (i.e., dipole aligned in the direction of the field) to pE for $\theta = 180^\circ$ (i.e., dipole in the direction opposite to the field).

Therefore, the preferred orientation of a dipole, with the minimum potential energy, is in the

direction of the field. The tendency of the dipoles to align in the direction of the field is crucial for understanding the behaviour of dielectrics.

7.3 DIPOLE MOMENT AND POLARIZATION

When a dielectric slab is placed in an electric field, e.g., between the parallel plates of a capacitor (Fig. 7.3). The dipoles are created and they tend to align themselves parallel to the direction of the field. As a result of which the charges appear on the opposite surfaces of the slab; positive charges on the surface facing negative electrode and negative charges on the surface facing positive electrode. These charges are the induced charges and the dielectric on the whole continues to be neutral. In this condition, the dielectric is said to be *polarized* and the phenomenon is termed *polarization*. Polarization P is defined as *dipole moment per unit volume* and is expressed as

$$P = Np \quad \dots(7.6)$$

or

$$P = \sum_{i=1}^N p_i$$

where N is the number of dipoles per unit volume (i.e., dipole concentration) and p is the dipole moment per unit volume. Here, it is being assumed that all the dipoles point in the same direction.

Example 7.1

In a water drop of radius 1 mm all the molecular dipoles point in the same direction. If the dipole moment of a water molecule is 6×10^{-30} cm, find the polarization in the water drop.

Solution.

Molecular mass of water = 18 gm

18 gm of water contain 6.023×10^{23} molecules.

$$\therefore \frac{18}{10^3} \text{ m}^3 \text{ of water will contain } 6.023 \times 10^{26} \text{ molecules.}$$

$$\text{Volume of water drop} = \frac{4\pi}{3} (10^{-3})^3 \text{ m}^3$$

$$\therefore \text{No of molecules in the drop} = \frac{6.023 \times 10^{26} \times 4\pi \times 10^{-9}}{18 \times 3} = 1.4 \times 10^{17} \text{ m}^{-3}$$

Polarization $P = Np$

Here, $p = 6 \times 10^{-30}$ cm, $N = 1.4 \times 10^{17} \text{ m}^{-3}$

$$\therefore P = 1.4 \times 10^{17} \times 6 \times 10^{-30} = 8.4 \times 10^{-13} \text{ cm}^{-2}$$

7.4 DIELECTRIC CONSTANT

In the polarized state, a dielectric is characterized by an excess charge due to the relative displacement of the charges within it, although the dielectric on the whole remains electrically neutral. So polarization is a phenomenon in which alignment of positive and negative charges takes place within the dielectric. When a medium is polarized, its electrostatic properties changes. In fact, polarization measures the additional flux density due to the presence of the (polarized) dielectric

material as compared to the free space. Therefore, it has the same unit as electric flux density \mathbf{D} (charge per unit area, C/m²). It is expressed as

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad \dots(7.7)$$

where \mathbf{D} is the electric displacement vector and \mathbf{E} is the electric field in the dielectric medium. The electric displacement vector \mathbf{D} depends only on the externally applied electric field and is not affected by the polarization of the medium. If \mathbf{E}_o is the externally applied electric field on the dielectric, then

$$\mathbf{D} = \epsilon_0 \mathbf{E}_o \quad \dots(7.8)$$

Using Eqs. (8.7) and (8.8), we have

$$\mathbf{E} = \mathbf{E}_o - \frac{\mathbf{P}}{\epsilon_0} \quad \dots(7.9)$$

This indicates that the field inside the dielectric is less than the externally applied electric field, i.e., the effect of polarization is to reduce the field inside the dielectric. The expression (7.7) is also usually written as

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \epsilon_r \mathbf{E} \quad \dots(7.10)$$

where, ϵ_0 is the absolute permittivity of the medium, $\epsilon_0 (= 8.854 \times 10^{-12}$ Farad/m) is the permittivity or dielectric constant of free space (vacuum) and ϵ_r is the relative dielectric constant or relative permittivity of the dielectric material (For free space $\epsilon_r = 1$, $\epsilon = \epsilon_0$). The relative dielectric constant (or simply dielectric constant)

$$\epsilon_r = \frac{\epsilon}{\epsilon_0} \quad \dots(7.11)$$

is an important parameter of the material. All the dielectric and optical characteristics of the material are contained in this constant. Therefore, by measuring the dielectric constant of a material we can gain an insight into its characteristics. We discuss below the procedure for measuring the relative dielectric constant of a material. Most of the times ϵ_r is simply referred to as dielectric constant since we rarely use the actual dielectric constant $\epsilon (= \epsilon_0 \epsilon_r)$.

To determine the dielectric constant we use a parallel plate capacitor. The plates of the capacitor are charged by connecting a battery, as shown in Fig. 7.3(a). When there is no electric field between the plates of the capacitor, the electric field E_o between the plates can be determined using the relation

$$E_o = \frac{V_o}{d} \quad \dots(7.12)$$

where V_o is the potential difference between the plates and d the separation between the plates. Now, the dielectric is inserted between the plates. The applied field E_o induces polarization in the dielectric, i.e., dipoles are aligned in the direction of the field. While the dipolar charge inside the dielectric cancels out, positive charges on the right and negatives charge on the left produces a field opposing the external field E_o (Fig. 7.3(b)). Let the new field be E . This can be determined by measuring the potential difference V across the plates by a voltmeter and using the relation.

$$E = \frac{V}{d} \quad \dots(7.13)$$

The dielectric constant ϵ_r can then be calculated in terms of the fields E_o and E by using the relation.

$$\epsilon_r = \frac{E_o}{E} \quad \dots(7.14)$$

This shows that the dielectric constant of a medium is always greater than unity as $E < E_o$. In terms of potential difference, the dielectric constant can be calculated using the relation.

$$\epsilon_r = \frac{V_o}{V} \quad \dots(7.15)$$

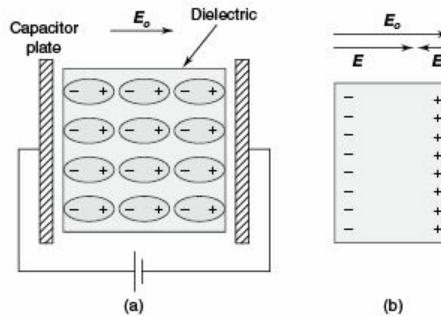


Fig. 7.3 (a) A dielectric inside field (b) a polarized dielectric.

Alternatively, the dielectric constant can be calculated by measuring the capacitance without the dielectric (or free space) between the plates $C_o = \frac{\epsilon_o A}{d}$ and with dielectric between the plates

$$C = \frac{\epsilon_o \epsilon_r A}{d} = \epsilon_r C_o \text{ and taking the ratio}$$

$$\epsilon_r = \frac{C}{C_o} \quad \dots(7.16)$$

Example 7.2

A barium titanate crystal is inserted in a parallel plate capacitor of area $15 \text{ mm} \times 10^{-9} \text{ F}$, calculate the relative dielectric constant of the crystal ($\epsilon_o = 8.854 \times 10^{-12} \text{ F m}^{-1}$).

Solution.

We know,

$$C = \epsilon_o \epsilon_r \frac{A}{d}$$

Here,

$$C = 10.5 \times 10^{-9} \text{ F}, A = 15 \times 10^{-3} \text{ m} \times 10^{-2} = 15 \times 10^{-5} \text{ m}^2$$

$$d = 2 \times 10^{-3} \text{ m}$$

$$\text{Relative dielectric constant } \epsilon_r = \frac{Cd}{\epsilon_0 A} = \frac{1.5 \times 10^{-9} \times 2 \times 10^{-3}}{8854 \times 10^{-12} \times 15 \times 10^{-5}} \\ = 2258$$

7.5 DIELECTRICS AND COULOMB'S LAW

In a capacitor with free space between the plates, the force between the two charges on the plates is proportional to the field E_o between the plates, i.e.,

$$F \propto E_o$$

After the introduction of the dielectric, this force reduces as ($E < E_o$).

$$F \propto E$$

Therefore,

$$\frac{F}{F_o} = \frac{E}{E_o} = \frac{1}{\epsilon_r} \quad \dots(7.17)$$

or, $F = \frac{F_o}{\epsilon_r}$...(7.18)

Thus, *on the introduction the dielectric between the plates, the force between the charges reduces by $\frac{1}{\epsilon_r}$.*

The Coulomb's law in free space is

$$F = \frac{1}{4\pi\epsilon_0} \cdot \frac{q_1 q_2}{r_{12}^2}$$

which in the presence of dielectrics takes the form

$$\begin{aligned} F &= \frac{1}{4\pi\epsilon} \cdot \frac{q_1 q_2}{r_{12}^2} \\ &= \frac{1}{4\pi\epsilon_0\epsilon_r} \cdot \frac{q_1 q_2}{r_{12}^2} \end{aligned} \quad \dots(7.19)$$

where $\epsilon_r = \frac{\epsilon}{\epsilon_0}$ is the relative permittivity or the dielectric constant of the medium.

7.6 GAUSS'S LAW IN DIELECTRICS

We know that Gauss's law is applied to parallel plate capacitor in free space, i.e., in absence of dielectric medium, at any point in the Gaussian surface is given by

$$\int E_o \cdot dA = \frac{q_o}{\epsilon_0}$$

or $E_o A = \frac{q_o}{\epsilon_0}$

or

$$E_o = \frac{q_o}{A \epsilon_o} \quad \dots(7.20)$$

where q_o is the charge enclosed by the Gaussian surface and A the area of plates. In the presence of dielectric between the plates (Fig. 7.40). The charge enclosed by the Gaussian surface is $q_o - q'$, where q' is the induced surface charge, also called the bound charges. If E is field now, then

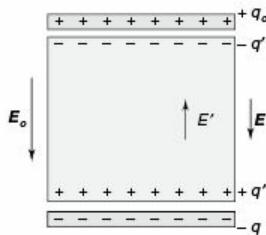


Fig. 7.4 A parallel plate capacitor with a dielectric between the plates.

$$\int \mathbf{E} \cdot d\mathbf{A} = \frac{q_o - q'}{\epsilon_o} \quad \dots(7.21)$$

or

$$EA = \frac{q_o}{\epsilon_o} - \frac{q'}{\epsilon_o}$$

or

$$E = \frac{q_o}{A \epsilon_o} - \frac{q'}{A \epsilon_o} \quad \dots(7.22)$$

The above expression indicates that the induced charge q' decreases the strength of the electric field between the plates. Using $E = \frac{E_o}{\epsilon_r}$, the above expression can be put as

$$\frac{E_o}{\epsilon_r} = \frac{q_o}{A \epsilon_o} - \frac{q'}{A \epsilon_o}$$

or

$$\frac{q_o}{A \epsilon_o \epsilon_r} = \frac{q_o}{A \epsilon_o} - \frac{q'}{A \epsilon_o} \quad \dots(7.23)$$

or

$$q_o - q' = \frac{q_o}{\epsilon_r} \quad \dots(7.24)$$

This indicates that the magnitude of induced charge q' is less than the charge on the plates q_o , also called free charge.

Further, if $\epsilon_r = 1$, $q' = 0$, i.e., in the absence of dielectric there is no induced charge.

Putting Eq. (7.2) in Eq. (7.21), we have

$$\int \mathbf{E} \cdot d\mathbf{A} = \frac{q_o}{\epsilon_o \epsilon_r}$$

or

$$\epsilon_r \int \mathbf{E} \cdot d\mathbf{A} = \frac{q_o}{\epsilon_o} \quad \dots(7.25)$$

This is the Gauss's law in the presence of dielectric. Thus, the effect of the polarization of the dielectric on the Gauss's theorem is to multiply the electric field by ϵ_r .

Equation (7.23) can also be rewritten as

$$\frac{q_o}{A} = \epsilon_o \left(\frac{q_o}{A \epsilon_o \epsilon_r} \right) + \frac{q'}{A} \quad \dots(7.26)$$

The quantity in bracket is the electric field \mathbf{E} in the dielectric. While the term $\frac{q'}{A}$ is the induced surface charge per unit area, known as electric polarization P or $P = \frac{q'}{A}$.

7.7 ELECTRICAL SUSCEPTIBILITY AND POLARIZABILITY

Equation (7.26) can also be put as

$$\begin{aligned} \frac{q_o}{A \epsilon_o} &= \mathbf{E} + \frac{q'}{A \epsilon_o} \\ \text{or} \quad \frac{q_o}{A \epsilon_o} &= \mathbf{E} + \frac{\mathbf{P}}{\epsilon_o} \\ \text{or} \quad \frac{q_o}{A} &= \epsilon_o \mathbf{E} + \mathbf{P} \end{aligned} \quad \dots(7.27)$$

The quantity on the right-hand side occurs quite often and is known as electric displacement vector \mathbf{D} . Therefore,

$$\mathbf{D} = \epsilon_o \mathbf{E} + \mathbf{P} \quad \dots(7.28)$$

And $\mathbf{D} = \frac{q_o}{A} = \frac{q_o}{A \epsilon_o} \times \epsilon_o$

or $\mathbf{D} = \epsilon_o \mathbf{E}_o = \epsilon_o \epsilon_r \mathbf{E} \quad \dots(7.29)$

or $\mathbf{D} = \epsilon \mathbf{E}$

$$\begin{aligned} \text{Also, polarization } \mathbf{P} &= \frac{q'}{A} = \frac{q_o}{A} \left(1 - \frac{1}{\epsilon_r} \right) \\ &= \epsilon_o \epsilon_r \mathbf{E} - \epsilon_o \mathbf{E} \\ &= \epsilon_o (\epsilon_r - 1) \mathbf{E} \end{aligned} \quad \dots(7.30)$$

$$\text{or} \quad \frac{P}{\epsilon_0 E} = \epsilon_r - 1 = \chi \quad \dots(7.31)$$

where χ is called the electric susceptibility of the dielectric medium. It is clear from Eq. (8.31) that

$$\epsilon_r = 1 + \chi \quad \dots(7.32)$$

Thus, the departure of the dielectric constant from unity, the value for vacuum, is equal to the electric susceptibility. (This is strictly true for isotropic medium but most of the actual dielectric media are anisotropic and therefore the value of ϵ_r or χ depends on the direction of the field, i.e., ϵ_r and χ are tensor quantities. However, to keep things simple, we assume the dielectric to be an isotropic medium in which case the dielectric constant is a scalar or a single number.)

It is quite reasonable to believe that polarization, i.e., generation and alignment of the dipoles will be dependent on the magnitude of the electric field. The molecular dipole moment can therefore be represented as

$$p \propto E$$

$$\text{or} \quad p = \alpha E \quad \dots(7.33)$$

where the constant α is called the *polarizability* of the molecule. Therefore, the polarization P can be written as

$$P = Np = N \alpha E \quad \dots(7.34)$$

where N is the number of dipoles per unit volume.

Substituting this in Eq. (8.28), we have

$$\begin{aligned} D &= \epsilon_0 E + N \alpha E \\ &= \epsilon_0 \left(1 + \frac{N \alpha}{\epsilon_0} \right) E \end{aligned} \quad \dots(7.35)$$

Comparing this with $D = \epsilon_r \epsilon_0 E$, we get

$$\epsilon_r = 1 + \frac{N \alpha}{\epsilon_0} \quad \dots(7.36)$$

This expression relates dielectric constant, a *macroscopic quantity*, to polarizability, a *microscopic quantity*.

Also, comparing Eq. (7.36) with (7.32), we have

$$\chi = \frac{N \alpha}{\epsilon_0} \quad \dots(7.37)$$

which reveals the dependence of χ on N and α .

In terms of the density of medium P

$$N = \frac{\rho N_A}{M} \quad (N_A = \text{Avogadro's number and } M = \text{molar mass})$$

We can express Eq. (7.36) as

$$\epsilon_r = 1 + \left(\frac{\rho N_A}{\epsilon_0 M} \right) \alpha \quad \dots(7.38)$$

This indicates that ϵ_r is directly proportional to the density of the dielectric medium. The experimental results show that Eq. (7.38), though holds for gases, it is not strictly true for liquids and solids. It is because the effective field in polarizing the dielectric is not E as assumed but different from E and, in fact, Eq. (7.33) should be represented.

$$P = \alpha E_{\text{local}} \quad \dots(7.39)$$

where E_{local} is the effective polarizing field called the *local electric field* or simply *local field*.

Example 7.3

The relative dielectric constant of polystyrene is 2.5. Find the polarization produced when a 0.5 mm thick sheet of polystyrene is subjected to 220 V.

Solution.

We know,

$$P = \epsilon_0 (\epsilon_r - 1) E$$

Here,

$$\epsilon_r = 2.5$$

$$E = \frac{V}{d} = \frac{220}{0.5 \times 10^{-3}} = 4.4 \times 10^5 \text{ V/m}$$

and

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ C/V.m}$$

$$P = 8.85 \times 10^{-12} \times (2.5 - 1) \times 4.4 \times 10^5 \text{ C/m}^2 = 5.84 \times 10^{-6} \frac{\text{C}}{\text{m}}$$

Example 7.4

The relative dielectric constant of a glass and plastic are 3.9 and 2.1 respectively. What voltage should be applied between electrodes that are separated with 1 mm of glass if we want the same charge density as would develop on another set of electrodes separated by 0.42 cm of plastic and 220 V?

Solution.

We know, charge density $D = \epsilon_0 \epsilon_r E$

$$\text{With plastic: } D = 8.85 \times 10^{-12} \times 2.1 \times \frac{220}{0.42 \times 10^{-2}}$$

$$\text{With glass: } D = 8.85 \times 10^{-12} \times 3.9 \times \frac{V}{1 \times 10^{-3}}$$

$$\therefore \frac{3.9V}{10^{-3}} = \frac{2.1 \times 220}{0.42 \times 10^{-2}}$$

$$\therefore V = \frac{2.1 \times 220 \times 10^{-3}}{3.9 \times 0.42 \times 10^{-2}} = 28.2 \text{ volts}$$

Example 7.5

A potential difference of 2.25 V is applied across the plates 1 cm \times 2 cm each of a parallel plate capacitor. (a) How far apart must the plates be to produce a charge density of $8.0 \times 10^{-6} \text{ C/m}^2$? (b) How many electrons accumulate on the negative plate?

Solution.

(a) Charge density $D = \epsilon_0 E$

$$\text{Here, } 8.0 \times 10^{-6} \frac{\text{C}}{\text{m}^2} = 8.85 \times 10^{-12} \frac{\text{C}}{\text{V.m}} \times \frac{2.25\text{V}}{d\text{m}}$$

\therefore Separation between the plates

$$d = \frac{8.85 \times 10^{-12} \times 2.25}{8.0 \times 10^{-6}}$$

$$= 2.49 \times 10^{-6} \text{ m} = 0.00249 \text{ mm}$$

$$\begin{aligned} \text{(b) Number of electrons} &= \frac{\text{Charge}}{\text{Charge on an electron}} = \frac{\text{Charge density} \times \text{area}}{\text{Charge on an electron}} \\ &= \frac{8.85 \times 10^{-12} \times 2.25 \times 10^{-2} \times 2 \times 10^{-2}}{1.6 \times 10^{-19} \times 2.49 \times 10^{-6}} \\ &= 9.99 \times 10^9 \text{ electrons} \end{aligned}$$

Example 7.6

The space between the plates of a parallel plate capacitor of capacitance $4.7 \mu\text{F}$ is filled with a dielectric of relative permittivity $\epsilon_r = 250$. (a) If a potential difference of 500 volt is applied across the capacitor, calculate the energy stored in the capacitor and the dielectric. (b) If E is the field strength applied to the dielectric, show that the energy stored per polarizing atom is $\frac{1}{2} \alpha E^2$, where α is the polarizability of the dielectric.

Solution.

(a) Energy stored in a capacitor, $W = \frac{1}{2} C V^2$

Here, $C = 4.7 \mu F = 4.7 \times 10^{-6} F, V = 500 \text{ volt}$

$$W = \frac{1}{2} \times 4.7 \times 10^{-6} \times 500 \times 500 = 0.587 \text{ Joule}$$

Capacitance of capacitor without dielectric

$$C_o = \frac{C}{\epsilon_r} = \frac{4.7}{250} \mu F = 0.019 \mu F$$

\therefore Energy without dielectric

$$\begin{aligned} W_o &= \frac{1}{2} C_o V^2 = \frac{1}{2} \times 0.019 \times 10^{-6} \times 500 \times 500 \\ &= 0.00237 \text{ Joule} \end{aligned}$$

\therefore Energy stored by the dielectric $W_d = W - W_o = 0.587 - 0.00237 = 0.5846 \text{ J}$

(b) As seen above, the energy stored in the polarizing dielectric is

$$\begin{aligned} W_d &= \frac{1}{2} (C - C_o) V^2 \\ &= \frac{1}{2} \left[\frac{\epsilon_o \epsilon_r A}{d} - \frac{\epsilon_o A}{d} \right] V^2 = \frac{A}{2} \frac{V^2}{d} (\epsilon_r - 1) \end{aligned}$$

Also,

$$P = \epsilon_o E (\epsilon_r - 1) \Rightarrow (\epsilon_r - 1) = \frac{P}{\epsilon_o E}$$

$$\therefore W_d = \frac{A \epsilon_o V^2}{2d} \cdot \frac{P}{\epsilon_o E} = \frac{AV^2}{2d} \cdot \frac{P}{E}$$

\therefore Energy stored per unit volume

$$\begin{aligned} \frac{W_d}{\text{volume}} &= \frac{W_d}{Ad} = \frac{1}{Ad} \frac{AV^2}{2d} \cdot \frac{P}{E} \\ &= \frac{1}{2} \frac{V^2}{d^2} \cdot \frac{P}{E} = \frac{1}{2} \cdot \frac{PE^2}{E} \quad \left(\because \frac{V}{d} = E \right) \\ &= \frac{1}{2} PE = \frac{1}{2} N \propto E^2 \quad (\therefore P \propto NE) \end{aligned}$$

\therefore Energy stored per polarizing atom

$$\frac{W_d}{N \text{ volume}} = \frac{1}{2} \propto E^2$$

7.8 LOCAL ELECTRIC FIELD: CLAUSSIUS-MOSSOTTI EQUATION

The effective electric field at an atomic site of a dielectric placed in externally applied field, called local field, is significantly different from the applied electric field. The effective electric field at the location of an electric dipole can only be equal to external field if no other dipoles are present in its

vicinity. In fact, in a solid with many other dipoles around, the effective field is also affected by the fields produced by other dipoles at the site under consideration.

To evaluate local field E_{loc} , we need to calculate the total field acting at the dipole under consideration, both due to the external field and the field due to all other dipoles in the dielectric. This was done by Lorentz, as follows. Consider a dipole at the centre of a (fictitious) spherical cavity whose radius r is sufficiently large so that the medium outside the cavity can be treated as a continuous medium as far as the dipole is concerned (Fig. 7.5). The interaction of the dipole with the other dipoles lying inside the cavity is however, to be treated microscopically because the discrete nature of the medium very close to the dipole should be taken into account. The local field at the dipole is written as

$$E_{\text{loc}} = E_o + E_1 + E_2 + E_3 \quad \dots(7.40)$$

where, E_o is the externally applied field due to the fixed charges external to the dielectric specimen,

E_1 is the field due to the polarization of the dielectric, i.e., induce surface charges on the dielectric specimen,

E_2 is the field (also called Lorentz field) due to the polarization charges lying on the surface of cavity (or Lorentz sphere) and

E_3 is the field due to other dipoles within the cavity.

Calculation of E_1 : The value of E_1 , also called *depolarization field* (because this field is in a direction opposite to external field), depends on the shape of dielectric specimen. For a dielectric specimen in the form of a infinite slab, using Gauss's law, it is found that

$$E_1 = - \frac{\mathbf{P}}{\epsilon_o} \quad \dots(7.41)$$

Customarily, the macroscopic field is expressed as

$$\mathbf{E} = \mathbf{E}_o + \mathbf{E}_1$$

Hence, E_{local} is also expressed as

$$E_{\text{loc}} = E + E_2 + E_3$$

Calculation of E_2 : As shown in Fig. 7.5(b), we consider the Lorentz cavity to be spherical in shape for simplicity. For regions external to the cavity, the distances to the cavity centre are so large (as cavity is supposed to be large) that we can treat the dipole distribution as continuous, resulting in macroscopic polarization \mathbf{P} . The polarization produces a field that can be described in terms of polarization charges on the surface of the cavity. The charge density due to the polarization of the imaginary cavity is written as

$$\sigma = - P_n \quad \dots(7.42)$$

where, subscript n denotes the normal component of \mathbf{P} . Therefore, the charge on circular element at polar angle θ is given by

$$d_q = - P \cos \theta \cdot 2\pi r \sin \theta \cdot n d\theta$$

The field produced by this charge at the centre of cavity, by Coulomb's law, is given by

$$dE_2 = - \frac{1}{4\pi\epsilon_o} \cdot \frac{d_q}{r^2} \cos \theta$$

Therefore, the field due to the total charge on the surface of the cavity is

$$\begin{aligned}
 E_2 &= \int_0^\pi \left(-\frac{P \cos \theta}{4\pi\epsilon_0 r^2} \right) \cos \theta (2\pi r^2 \sin \theta d\theta) \\
 &= \frac{P}{2\epsilon_0} \int_0^\pi \cos^2 \theta \sin \theta d\theta \\
 &= \frac{1}{3\epsilon_0} P
 \end{aligned} \quad \dots(7.43)$$

The direction of this field is the same as that of E_o .

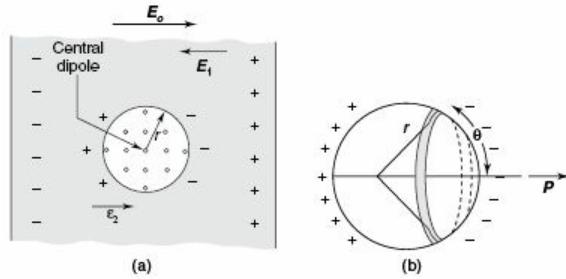


Fig. 7.5 (a) A polarized dielectric (b) Cavity.

Calculation of E_3 : This field, which is due to other dipoles in the cavity, can be evaluated by the fields due to the individual dipoles using Eq. (7.2). The resulting field depends on the crystal structure of the dielectric. Let us assume the dielectric crystal to have cubic symmetry (e.g., NaCl and CsCl). If the external field is applied along the z-direction, then in accordance with Eq. (7.2), the z-component of the field at the centre of the cavity due to all other dipoles within the cavity can be expressed as

$$E_{3z} = \frac{p}{4\pi\epsilon_0} \sum_i \frac{3Z_i^2 - r_i^2}{r_i^5}$$

where p is the magnitude of the dipole moment of a single dipole and it is assumed that all dipoles are parallel to the z-axis.

Due to cubic symmetry of the crystal, the x, y and z directions are equivalent, thus

$$\sum_i \frac{x_i^2}{r_i^5} = \sum_i \frac{y_i^2}{r_i^5} = \sum_i \frac{z_i^2}{r_i^5}$$

Therefore, for a crystal with cubic symmetry $E_3 \equiv E_{\text{dip}} = 0$.

Substituting the values of E_1 , E_2 and E_3 in Eq. (7.40), we get

$$E_{\text{loc}} = E_o - \frac{2}{3} P \quad \dots(7.44)$$

which gives the effective polarizing field in terms of the applied electric field and the polarization. Comparing the value of E_{local} obtained above with that of E in Eq. (7.9), we find

$$E_{\text{local}} = E + \frac{1}{3\epsilon_0} P \quad \dots(7.45)$$

This is known as *Lorentz relation*. This expression shows that $E_{\text{local}} > E$, i.e., the molecules are more effectively polarized than what it would have been by E alone. The difference between E , also known as Maxwell field, and the Lorentz field E_{loc} can be understood with the help of the following explanation. The field E is an average macroscopic field, the average being taken over a large number of molecules (Fig. 7.6) This field is used in Maxwell equations for the macroscopic description of dielectric media. In the present situation the field E is a constant throughout the media. On the other hand, E_{loc} is a microscopic field which varies within the medium. As indicated in the figure, this field is quite large at the molecular sites and hence the molecules are more effectively polarized than they would be in the average field E .

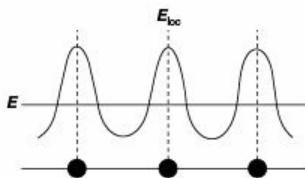


Fig. 7.6 Maxwell field E and the local field E_{loc} , the solid circles represent molecules.

Claussius-Mossotti relation. Using Eqs. (7.34) and (7.39) the polarization can be expressed as

$$P = N p = N \alpha E_{\text{local}} \quad \dots(7.46)$$

where α is the polarizability of the considered atom or molecule.

Substituting the value of E_{local} from Eq. (7.45), we get

$$\begin{aligned} P &= N \alpha \left(E + \frac{1}{3\epsilon_0} P \right) \\ \text{or} \quad \left(1 - \frac{N \alpha}{3\epsilon_0} \right) P &= N \alpha E \\ \text{or} \quad P &= \left(\frac{N \alpha}{1 - \frac{N \alpha}{3\epsilon_0}} \right) E \end{aligned} \quad \dots(7.47)$$

The denominator, in the above expression, being less than unity, it is evident that polarization is higher than what expected from Eq. (7.34); the enhancement is due to the local field correction. Substituting Eq. (7.47) in Eq. (7.36), we get

$$\epsilon_r = 1 + \frac{1}{\epsilon_a} \left(\frac{N \alpha}{1 - \frac{N \alpha}{3 \epsilon_a}} \right)$$

which on simplification yields

$$\epsilon_r = \frac{1 + \frac{2}{3} \frac{N \alpha}{\epsilon_a}}{1 - \frac{N \alpha}{3 \epsilon_a}} \quad \dots(7.48)$$

This is generalization of the expression for ϵ_r when the local field correction is taken into account.

In gases, the molecular concentration N is small, therefore, $\frac{N \alpha}{3 \epsilon_a} \ll 1$. Expanding the denominator

in Eq. (7.48) in powers of $\frac{N \alpha}{3 \epsilon_a}$, and simplifying in terms of first order, we get the same expression as given in Eq. (7.36).

This can be understood as for small N the polarization P is also small, which in accordance with Eq. (7.43) implies that the local field is not much different from the average field. This is not so in liquids and solids and Eq. (7.48) has wide applicability. Eq. (7.48) can also be written as

$$\begin{aligned} \epsilon_r - \frac{N \alpha \epsilon_r}{3 \epsilon_a} &= 1 + \frac{2N \alpha}{3 \epsilon_a} \\ \text{or} \quad \epsilon_r - 1 &= \frac{N \alpha}{3 \epsilon_a} (\epsilon_r + 2) \\ \text{or} \quad \frac{\epsilon_r - 1}{\epsilon_r + 2} &= \frac{N \alpha}{3 \epsilon_a} \end{aligned} \quad \dots(7.49)$$

This equation is known as *Claussius-Mossotti relation* and relates relative permittivity or dielectric constant, a macroscopically measurable physical quantity, to polarizability, a molecular microscopic quantity.

Multiplying Eq. (7.49) by molar volume $\frac{M}{\rho}$ (M = molecular mass, ρ = density), we get

$$\begin{aligned} \frac{\epsilon_r - 1}{\epsilon_r + 2} \frac{M}{\rho} &= \frac{M \alpha}{3 \epsilon_a} \cdot \frac{M}{\rho} \\ \text{As} \quad \frac{NM}{\rho} &= N_A \\ \therefore \quad \frac{\epsilon_r - 1}{\epsilon_r + 2} \frac{M}{\rho} &= \frac{N_A \alpha}{3 \epsilon_a} = p_m \end{aligned} \quad \dots(7.50)$$

where p_m is molar polarization and α is the molar polarizability which can be calculated by knowing M (molar mass), ρ (density) and ϵ_r (dielectric constant).

7.9 TYPES OF POLARIZATION

The total polarization observed in a specimen may comprise the following four components: (i) electronic polarization P_e , (ii) orientational (or dipolar) polarization P_o , (iii) ionic polarization P_i and (iv) space charge polarization P_s . Therefore, the total polarization P may be expressed as

$$P = P_e + P_o + P_i + P_s \quad \dots(7.51)$$

In terms of polarizability, the total polarizability α of a specimen can be expressed as

$$\alpha = \alpha_e + \alpha_o + \alpha_i + \alpha_s \quad \dots(7.52)$$

where α_e , α_o , α_i and α_s are the electronic, orientational, ionic and space-charge polarizability respectively.

Electronic Polarization: Under the action of an electric field, the centres of the electrons and the nuclei in the atoms of the molecules in the specimen get displaced in opposite directions (Fig. 7.7). This process of inducing dipoles is called electronic polarization. This ratio of induced dipole moment to the local field is called *electronic polarizability*, α_e . The electronic polarizability arises even in the neutral atom due to the relative displacement of orbital electrons. Electronic polarizability is independent of temperature.

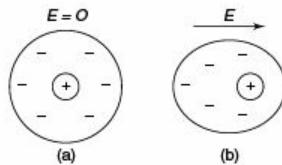


Fig. 7.7 Electronic polarization (a) unpolarized atom (b) polarized atom.

Orientational (or dipolar) Polarization: If the molecules of a dielectric specimen have permanent moment, i.e., dipole moment even in the absence of an externally applied electric field, the molecule is said to be a dipolar molecule and the specimen, a dipolar substance. For example, a water (H_2O) molecule is a dipolar molecule (Fig. 7.8 (a)). The dipole moments of the two OH bonds add vectorially to give a net dipole moment. While in the case of CO_2 molecule (Fig. 7.8(b)), the two CO bonds cancel each other because of the rectilinear shape of the molecule resulting in a zero net dipole moment. Thus, whether a molecule possesses a permanent moment or not depends on its geometrical arrangement of charges.

Even though the individual molecules of a dipolar solid possess permanent dipole moment, in the absence of an externally applied electric field, the moments are randomly distributed and the net polarization is zero. However, when an electric field is applied, the molecular dipoles tend to align in the direction of the field resulting in a net polarization. *This polarization is called the orientational (or dipolar) polarization due to the fact that it is because of the orientation of the pre-existing dipoles.* On increasing the temperature of the specimen, due to the increased energy of the dipoles, the alignment reduces. The corresponding polarizability is called *orientational polarizability* α_o .

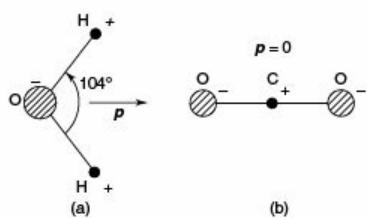


Fig. 7.8 (a) The water molecule having permanent dipole moment. (b) CO₂ molecule with no dipole moment.

Ionic Polarization: If the molecules of the specimen contain ionic bond (e.g., NaCl), the positively charged ions (Na⁺) and the negatively charged ions (Cl⁻) are displaced in opposite directions (Fig. 7.9) resulting in an increase in the bond length compared to the normal bond length. The effect of this increase in bond length is to produce a net dipole moment. The specimen gets polarized due to displacement of the ions and hence, this type of polarization is called *ionic polarization* and the corresponding polarizability is called *ionic polarizability* α_i . Ionic polarizability is also independent of temperature.

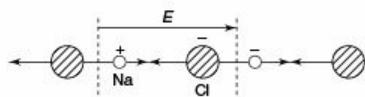


Fig. 7.9 Ionic polarization in NaCl.

Space-charge polarization: If the specimen is a multiphasic material, then the charges accumulate at the interfaces of the phases and at the electrodes. The ions diffuse over appreciable distances in response of an applied field leading to redistribution of charges in the dielectric medium. This leads to space-charge polarization and the corresponding polarizability is called *space charge polarizability* α_s .

The electronic polarization is present in all types of substances, but the presence of the other three types of polarization depends on the nature of the specimen. The ionic polarization is present in only ionic substances while in a dipolar substance, in addition to dipolar polarization, the ionic and electronic polarizations are also present. The space-charge polarization is present in only multiphasic substances.

The various polarizabilities can be segregated from each other because each component has its own characteristics which distinguish it from others. Orientational polarizability exhibits strong dependence on both frequency and temperature while the electronic and ionic components are nearly independent of temperature and frequency.

Figure 7.10 shows typical dependence of polarizability over a wide frequency range, from the static to the ultraviolet region. The space charge polarizability is present at the machine frequencies only (\approx 50 to 60 Hz, although it can extend up to 1 kHz in some cases). This component is the slowest process as it involves diffusion of ions over several interatomic separations. The relaxation time for this process is related to the frequency of ion jumps under the influence of the applied field, a typical value being 10^2 Hz, ω_s . The polarizability (Fig. 7.10) shows an expected decrease with the increase in frequency, as the successive polarizability component is damped out.

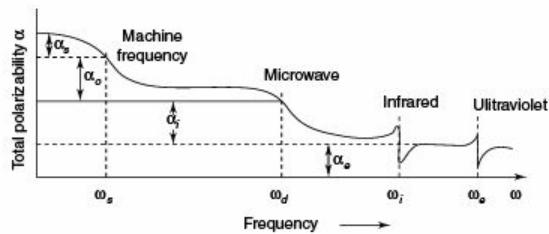


Fig. 7.10 Total polarizability α versus frequency ω .

The polarizability is nearly constant till a frequency ($\omega_d \approx 10^8$ Hz) is reached where the contribution from orientational polarizability also ceases and the polarizability decreases substantially. The contribution from this component is effectively present in the audio frequency range. The decrease in polarizability due to the ceasing of the dipolar contribution can be understood from the fact that in the frequency range $\omega > \omega_d$, *The applied field oscillates too rapidly for the dipoles to follow and so the dipoles remain essentially stationary.*

In between the frequency range ω_d and ω_i , the polarizability again remains constant and then plummets at the higher frequency. The frequency ω_i lies in the infrared region ($\sim 10^{13}$ Hz). The ionic polarization involves the displacement of heavier ions. The frequency with which ions can be displaced over a small fraction of the interatomic separation is of the same order as the lattice vibration frequency ($\sim 10^{13}$ Hz). If the frequency of the applied electric field is larger than this field, the heavy ions are no longer able to follow the rapidly oscillating field and consequently the ionic polarizability α_i vanishes.

In the frequency range above the infrared region, only the electronic polarizability remains effective as the electrons being very light, are able to follow the field of even higher frequencies, including the visible and ultraviolet regions. However, at still higher frequencies. ($\omega > \omega_e \approx 10^{15}$ Hz, beginning of ultraviolet region) even the electrons are unable to follow the rapidly oscillating field and even the electronic polarizability vanishes.

Thus, with frequencies, the dielectric constant of a substance decreases continuously and substantially as the frequency is increased from the static to the optical region. For example, in the static condition the dielectric constant of water is 81, which reduces to 1.8 at optical frequencies. The characteristic frequencies of dipolar ω_d and ionic ω_i polarizations depends on the nature of the specimen and therefore vary from material to material. However, the order of magnitudes remain in the region mentioned above, i.e., in the microwave and infrared regions respectively. Thus, by measuring polarizabilities in different frequency regions, we can determine various polarizability components which helps us understand the internal microscopic structure of the substance.

7.10 ELECTRONIC POLARIZATION

On the application of a *static electric field*, the centre of electronic shells and the centre of the nucleus do not coincide (Fig. 7.11). The electronic clouds shift towards the positive electrode while the nucleus, being massive, shifts to a smaller extent towards the negative electrode. Let the charge on the nucleus be Ze and the total electronic charge be $-Ze$ distributed over a sphere of radius. Thus, the charge density ρ is given by

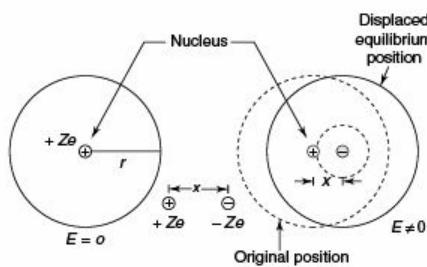


Fig. 7.11 (a) An atom without any field. (b) Electronic polarization.

$$\rho = \frac{Ze}{\frac{4}{3}\pi r^3}$$

If the applied electric field is E , the Lorentz forces experienced by the nucleus and electrons in the opposite is $Ze E$. Therefore, they are pulled apart in the opposite directions. As they are pulled apart, a Coulomb force develops between them which tends to counteract the displacement (of the order of 10^{-17} m at applied field ≈ 30 kV/m). Let this small displacement be x . We also assume that this displacement is in the electronic cloud only. Thus, Lorentz force $= -Ze E$

and Coulombic force $= Ze \times \left[\frac{\text{Charge enclosed in the sphere of radius } x}{4\pi\epsilon_0 x^2} \right] \quad \dots(7.53)$

$$\begin{aligned} \text{Charge enclosed in sphere of radius } x &= \frac{4}{3}\pi x^3 \cdot \rho \\ &= -\frac{4}{3}\pi x^3 = -\frac{Ze}{\frac{4}{3}\pi r^3} = -\frac{Ze x^3}{r^3} \end{aligned}$$

Substituting this in Eq. (7.53), we get

$$\begin{aligned} \text{Coulombic force} &= -Ze \cdot \frac{Ze x^3}{r^3} \cdot \frac{1}{4\pi\epsilon_0 x^2} \\ &= -\frac{Z^2 e^2 x}{4\pi\epsilon_0 r^3} \quad \dots(7.54) \end{aligned}$$

Under equilibrium, the two forces are equal.

$$\therefore -Ze E = \frac{Z^2 e^2 x}{4\pi\epsilon_0 r^3}$$

$$\text{or} \quad E = \frac{Ze x}{4\pi\epsilon_0 r^3}$$

$$\text{or } x = \frac{4\pi \epsilon_0 r^3}{Ze} E \quad \dots(7.55)$$

Thus, the displacement x is proportional to the applied field E . The resulting dipole moment is given by

$$\begin{aligned} \mu_e &= Ze x \\ \text{or } \mu_e &= \frac{Ze 4\pi \epsilon_0 r^3 E}{Ze} \\ \text{or } \mu_e &= 4\pi \epsilon_0 r^3 E \end{aligned}$$

$$\text{Also, } \mu_e = \alpha_e E$$

$$\therefore \text{Electronic polarizability } \alpha_e = 4\pi \epsilon_0 r^3 \quad \dots(7.56)$$

For $r = 10^{-10}$ m, $\alpha_e \approx 10^{-41}$ Farad. m². Thus, the electronic polarization, i.e., dipole moment per unit volume, is given by

$$P_e = N \mu_e = N \alpha_e E$$

where N is the number of atoms/m³.

$$\text{Also, } P_e = E \epsilon_0 (\epsilon_r - 1)$$

$$\therefore E \epsilon_0 (\epsilon_r - 1) = N \alpha_e E$$

$$\text{or } \epsilon_r - 1 = \frac{N \alpha_e}{\epsilon_0} \quad \dots(7.57)$$

Example 7.7

Calculate the polarizability of an argon atom if the relative permittivity of argon at NTP is 1.000435.

Solution.

$$\text{Number of atoms of argon at NTP} = \frac{6.023 \times 10^{26}}{22.4} = 2.69 \times 10^{25}$$

$$\text{We know, } \epsilon_0 (\epsilon_r - 1) = N \alpha_e$$

$$\begin{aligned} \alpha_e &= \frac{\epsilon_0 (\epsilon_r - 1)}{N} \\ \therefore \alpha_e &= \frac{8.85 \times 10^{-12} \times 0.000435}{2.69 \times 10^{25}} \\ &= 1.43 \times 10^{-40} \text{ Fm}^2 \end{aligned}$$

Example 7.8

An oxygen atom produced a dipole moment of 5.0×10^{-23} cm when subjected to an electric field. If the separation between the centre of electronic cloud and the nucleus is 4×10^{-17} m, find out the polarizability of the oxygen atom.

Solution.

If d is separation between the centres of negative and positive charges, then under equilibrium

$$(8e) E = \frac{(8e)(8e)}{4\pi \epsilon_0 d^2} \Rightarrow E = \frac{8e}{4\pi \epsilon_0 d^2}$$

Here, $d = 4 \times 10^{-17} \text{ m}$

$$\therefore E = \frac{8 \times 1.6 \times 10^{-19}}{4 \times 3.14 \times 8.85 \times 10^{-12} \times 16 \times 10^{-34}} = 2.6 \times 10^{24} \text{ Vm}^{-1}$$

Also, $p = \alpha E \Rightarrow \alpha = \frac{p}{E}$

$$\therefore \text{Polarizability } \alpha = \frac{5.0 \times 10^{-23}}{2.6 \times 10^{24}} = 1.9 \times 10^{-47} \text{ Fm}^2.$$

Example 7.9

The dielectric constant of helium measured at 0°C and one atmosphere is 1.000684; under this condition the gas contains 2.7×10^{25} atoms/m³. (a) Find out the radius of the electronic cloud (i.e., atomic radius). (b) What is the displacement of the cloud when a helium atom is subjected to an electric field of 10^6 Vm^{-1} .

Solution.

(a) We know, $\epsilon_0 (\epsilon_r - 1) = N \alpha_e$

$$\text{or } N 4\pi \epsilon_0 r^3 = \epsilon_0 (\epsilon_r - 1)$$

$$\text{or } r^3 = \frac{(\epsilon_r - 1)}{4\pi N}$$

$$\text{Here, } \epsilon_r - 1 = 0.000684$$

$$N = 2.7 \times 10^{25} \text{ atoms/m}^3$$

$$r^3 = \frac{0.000684}{4 \times 3.14 \times 2.7 \times 10^{25}}$$

$$\text{or } r = 5.87 \times 10^{-11} \text{ m}$$

$$(b) \text{ Displacement } x = \frac{4\pi \epsilon_0 E r^3}{Ze}$$

$$\text{Here, } E = 10^6 \text{ V/m}, \quad r = 5.87 \times 10^{-11} \text{ m}$$

$$Z = 2 \quad e = 1.6 \times 10^{-19} \text{ C}$$

$$\therefore x = \frac{4 \times 3.14 \times 8.85 \times 10^{12} \times 10^6 \times (5.87 \times 10^{-11})^3}{2 \times 1.6 \times 10^{-19}}$$

$$= 7.03 \times 10^{-17} \text{ m.}$$

7.11 DIELECTRIC DISPERSION: DIELECTRIC LOSS

Let us now discuss the behaviour of electric dipoles in an alternating electric field. In an oscillating electric field, the dipoles in the system tend to follow the variations in the field by flipping back and forth as the direction of the field reverses during each cycle. As long as the frequency of the a.c.

field remains low ($< 10^6$ Hz). The dipoles can follow the field variations. Therefore, the polarization follows the alternations of the field without any significant lag and the permittivity is almost independent of the frequency and has nearly the same value as in a d.c. field. However, with increasing frequency of the applied field, the dipoles are no longer able to rotate sufficiently rapidly and their oscillations begin to lag behind the field. At sufficiently higher frequency of the field, the dipoles are completely unable to follow the field and the contribution from the orientational polarization to the polarization of the dielectric ceases; This occurs in the *radio frequency range* ($10^6 - 10^{11}$ Hz). At still higher frequencies (*infrared region* $10^{11} - 10^{14}$ Hz), the relatively heavy positive and negative ions also fail to follow the field variations and contribution from ionic polarization also vanishes and only electronic polarization persists.

The above discussed phenomena lead to fall in the permittivity (or dielectric constant) of a dielectric material with increasing frequency and is often referred to as *anomalous dielectric dispersion*. Dispersion occurring due to the transition from maximum orientational polarization at zero or low frequencies to negligible orientational polarization at high radio frequencies is referred to as *dielectric relaxation*.

In response to an alternating field, the dipoles change their directions repeatedly and during this process undergo collision with other molecules in the substance. This results in the absorption of energy from the field, often referred to as *dielectric loss* or simply loss. Eventually this energy appears in the form of heat and raises the temperature of the substance. The dielectric loss at low frequencies is primarily due to d.c. resistivity. However, at higher frequencies, the dielectric loss is mainly due to dipolar rotation (or orientation), ionic transitions from the lower energy states to higher energy states, and electron hopping. Space charge polarization losses typically occur at low frequency (1 kHz) and are associated with ion migration and electrode contact losses or with the presence of grain boundaries or inhomogeneous phases in the dielectrics. Space charge losses in the latter case can occur up to much higher ($\sim 10^5$ Hz) frequencies. The dielectric losses falling in the infrared region are referred to as *infrared absorption* and that in the optical region are referred to as *optical absorption*.

7.12 DIELECTRIC SMART MATERIALS

We discuss below piezoelectric, ferroelectric and pyroelectric materials which are dielectric in nature and 'smart' in their behaviour. *The materials having the capability to sense their environment and respond accordingly are referred to as 'smart' or 'intelligent' materials.* These materials can 'sense' the change in their input (such as stress, heat, electric field, etc.) and correspondingly produce a changed output (such as charge, current, etc) and therefore can act as *sensor*. As these materials can also show converse effects they can also be used as *actuators*. Ferroelectric materials exhibit most of these functions and are often said to be *very smart materials*. Smart materials have immense application potentials.

Do you know that in the cellular phones which all of us carry nowadays, dielectric materials are used extensively? These materials are used as resonator (vibrator), piezoelectric receiver, piezoelectric speaker, chip monolithic ceramic capacitors, microwave oscillators, microwave filters, ceramic filters, etc.

7.13 PIEZOELECTRICITY

Certain materials produce electric charges on their surfaces when a mechanical stress is applied on them. The induced charges are proportional to the mechanical stress. This effect is called *piezoelectric*

effect and was discovered by Pierre and Jacques Curie in 1880. The meaning of the word *piezo* is pressure and piezoelectricity is also called pressure electricity as the application of a pressure on certain materials induces electricity in them. The induced charges on the surface produce an electric field inside the substance. The *inverse effect*, i.e., as the applied field produces a strain, has also been observed. Quartz is the most familiar and most widely used piezoelectric material. PZT (lead zirconate lead titanate) is another widely used material.

The piezoelectric effect is small in magnitude. An electric field of 1000 V/cm in quartz produces a strain of only 10^{-7} . That is, in a rod of length 1 cm, the change in length is only 10 \AA° . However, due to the inverse effect, even a small strain produces enormous electric field.

The piezoelectric effect is often used to convert electrical energy into mechanical energy and vice versa, i.e., the piezoelectric materials are used as *transducers*. These are also used extensively in the fabrication of devices such as actuators, surface acoustic wave (SAW) devices, frequency control, piezoelectric transformers, speakers and buzzers, ultrasonic generators, ultrasonic motors, etc.

The microscopic origin of piezoelectricity is the displacement of ionic charges within the crystal. In the absence of strain, the distribution of the charges at their lattice sites is symmetric and consequently the internal electric field is zero. But when a mechanical strain is applied, the charges are displaced. If the resulting charge distribution is not symmetric, a net polarization and an electric field develops in the substance. Naturally, piezoelectricity is shown by the substances, the *unit cell* of which *lacks a centre of symmetry*. As shown in Fig. 7.12(a), if a centre of inversion is present and it persists even after when a stress is applied, the net polarization remains zero and no piezoelectricity is shown by the substance. However, when there is no centre of inversion as shown in Fig. 7.12(b) for quartz (SiO_2), a distortion on application of a strain produces a net polarization and consequently the substance exhibits piezoelectricity.

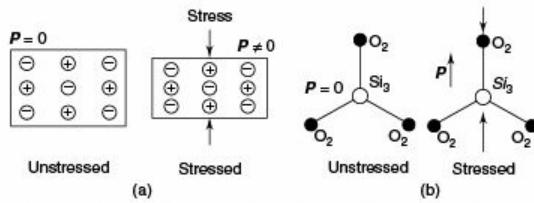


Fig. 7.12 (a) Crystal with centre of inversion exhibits no piezoelectric effect.
 (b) Origin of piezoelectric effect in quartz: it lacks a centre of inversion.

In crystallography, the elements of symmetry that are used to define symmetry about a point in space, for example, the central point of a unit cell, are: (i) a point (centre) of symmetry, (ii) axes of rotation, (iii) mirror planes and (iv) a combination of these. Using these symmetry elements, all crystals are divided into 32 different classes or point groups, as shown in Fig. 7.13. These 32 point groups are subdivided into seven basic crystal systems, which are in the order of ascending symmetry: triclinic, monoclinic, orthorhombic, tetragonal, rhombohedral (trigonal), hexagonal and cubic. Out of 32 point groups, 21 do not possess a centre of symmetry and 20 of these are piezoelectric, as one of these though lacking a centre of symmetry, is not piezoelectric because of the other combined symmetry elements. No regular cubic lattice can exhibit piezoelectricity. The sign of charge is dependent on the direction of the stress, whether compressive or tensile.

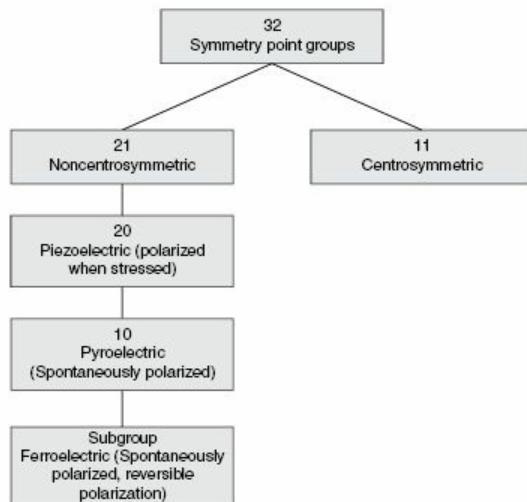


Fig. 7.13 Classification of piezoelectrics and subgroups on the basis of internal crystal structure.

Piezoelectric materials are used in delay lines in addition to transducers. When an electrical signal is converted into a mechanical wave, it travels through a quartz rod at the velocity of sound, which is much less than the speed of light, leads to considerable delay of the signal. Such delay lines are used in colour TV sets and videotape recorders. Also, piezoelectric and related electro-optic crystals are now widely used in the fields of laser technology and modern optics. For instance, the cavity length of a laser may be varied continuously in controlled manner by the application of a voltage to a piezoelectric crystal located at one end of the cavity. Piezoelectric transformers, because of their compact size in comparison to the conventional coil type transformers and so electromagnetic noise free, are being preferred in laptop computers, LCD TVs, etc.

Due to the lack of a centre of inversion and complex structure of piezoelectrics, piezoelectric effect or the converse effect (strain under electric field) is not isotropic, i.e., it is different in different directions or anisotropic. However, a simple representation of the phenomena in a schematic one-dimensional nation, the piezoelectric can be represented as

$$P = Zd + \epsilon_0 \chi E \quad \dots(7.58)$$

and the converse piezoelectric effect as

$$e = Z_s + Ed \quad \dots(7.59)$$

where P is the polarization, Z the stress, d the piezoelectric strain constant, E the electric field, χ the electric susceptibility, e the elastic strain and the s the elastic compliance constant. Eq. (7.58) represents the development of polarization by an applied stress and Eq. (7.59) the development of elastic strain by an applied electric field.

In practice, however, the tensile or compressional or shear strains may develop in different directions due to an applied field depending upon the crystal orientation and the field direction. Therefore, the *piezoelectric strain constants* are defined as

$$d_{ik} = \left(\frac{\partial e_k}{\partial E_i} \right)_z \quad \dots(7.60)$$

where $i = x, y, z$ and $k \equiv xx, yy, zz, xy, yz, zx$. Thus, it is third rank tensor.

7.14 PYROELECTRICITY

As indicated in Fig. 7.13, there are 10 crystal classes out of 20 piezoelectric crystals, which are referred to as *pyroelectric*. These crystals have the characteristic of being permanently polarized within a given temperature range. Unlike piezoelectric crystals, which get polarized when stressed, the pyroelectrics develop polarization spontaneously and form permanent dipoles in the structure. This *spontaneous polarization changes with temperature hence this phenomenon is referred to as pyroelectricity*. When the temperature of a pyroelectric crystal is changed, electric charges corresponding to the change of spontaneous polarization appear on the surface of the crystal. Pyroelectric crystals are also referred to as *electric stones*. It is observed that when such a stone is thrown in the fire, it generates electric charges causing a *cracking sound*. This is basically due to the temperature dependence of the spontaneous polarization of a polar material.

Pyroelectric materials have been used as temperature sensors and infrared light detectors.

Among the pyroelectric crystals, those whose spontaneous polarization can be reversed by an externally applied electric field are called ferroelectrics, which we discuss below in detail.

7.15 FERROELECTRICITY

The dielectric materials which exhibit spontaneous polarization, i.e., are polarized even in the absence of an applied electric field and whose polarization is reversible under suitable electric field are called *ferroelectric materials*, in analogy with ferromagnetic materials. The variation of polarization P with electric field E shows a hysteresis curve. The ferroelectric materials are said to be *very smart materials*. All ferroelectric materials are also pyroelectric and piezoelectric but the reverse is not true. Ferroelectricity, like piezoelectricity, can occur only in noncentrosymmetric crystals. In addition, for ferroelectricity, the crystal must also have a favourable axis of polarity. Only ten crystal classes have sufficiently low degrees of symmetry for the occurrence of ferroelectricity. Ferroelectric materials are highly useful as they can be used as high permittivity capacitors, ferroelectric memories, pyroelectric sensors, transducers, PTC (Positive Temperature Coefficient of resistance) thermistors, electro-optic devices, etc.

Every ferroelectric crystal is characterized by a temperature called ferroelectric curie temperature or transition temperature (T_c) above which it loses its ferroelectric behaviour and shows *paraelectric* behaviour (normal dielectric behaviour). A phase transition occurs at T_c , and as the crystal crosses over the curie point it changes from ferroelectric phase to paraelectric phase in which the dipoles of the crystal get randomly oriented.

The dielectric constant of a ferroelectric material changes with temperature according to the *Curie-Weiss law* given by

$$\epsilon_r = B + \frac{C}{T - T_c}, \quad T > T_c \quad \dots(7.61)$$

where B and C are constants independent of temperature. Constant C is called the *curie constant*. The above expression is represented by Fig. 7.14(a) and is valid for $T > T_c$. For the temperature $T < T_c$, the material becomes spontaneously polarized, i.e., an electric polarization develops in it without the help of an externally applied field. For $T > T_c$, the material is in paraelectric phase. Below T_c the dipoles interact with each other and form domains. All the dipoles within a domain point in a certain favourable direction. Figure 7.14(b) shows the variation of spontaneous polarization with temperature. For $T > T_c$, the interaction between the dipoles ceases, and the individual dipoles get randomly oriented. The material no longer possesses spontaneous polarization, i.e., it turns into paraelectric phase.

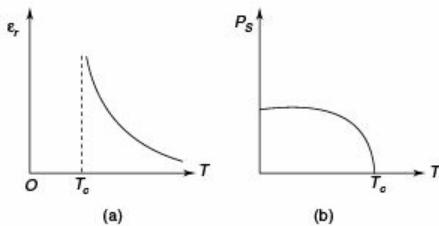


Fig. 7.14 (a) Variation of dielectric constant ϵ_r with temperature T for $T > T_c$.
 (b) Variation of spontaneous polarization with temperature.

In Eq. (7.61), the second term is much larger than the first one, e.g., if $\epsilon_r \approx 1000$, $B \approx 5$, so it is often a practice to ignore the first term and express the Curie-Weiss law as

$$\epsilon_r = \frac{C}{T - T_c} \quad \dots(7.62)$$

7.15.1 Ferroelectric Hysteresis Loop

When an electric field is applied to a ferroelectric specimen, the polarization first rises rapidly with the increasing applied field and then shows down and finally saturates. The linear extrapolation to zero field gives P_s , the *saturation* or *spontaneous polarization*. As the field is lowered, finally reducing to zero, the left-over polarization at zero field is called the *residual* or *remanent polarization* P_r . The electric field need to be applied in the opposite direction to reduce the polarization to zero, i.e., to depolarize the specimen, is called the *coercive field*, E_c . On further application of the field in the opposite direction the saturation polarization is obtained in the opposite direction. On again changing the direction of field the hysteresis curve (P - E loop) is obtained. A typical ferroelectric hysteresis loop is shown in Fig. 7.15. The exact hysteresis loop of a specimen and hence the parameter P_s , P_r and E_c depends upon a number of factor such as the dimension of specimen, the temperature, the microstructure, the thermal and electrical properties. The nature of the loop changes with the temperature specimen and for $T \geq T_c$ the loop disappears as the specimen is no longer ferroelectric.

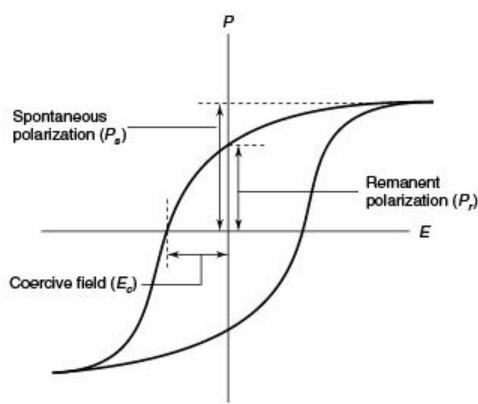


Fig. 7.15 Ferroelectric hysteresis loop.

In the absence of any external field, as the specimen is cooled below the curie temperature from the paraelectric phase, the polarization is nucleated at several points in the crystal. These nuclei normally differ in the direction of the polarization. As the nuclei grow through the crystal in the ferroelectric phase, they form regions called domains, differing in the direction of polarization. When an electric field is applied, the number and size of the domains polarized in the direction of the applied field increase. The extent of polarization increases with the increasing polarization and subsequently saturation is obtained. The reversal of the direction of polarization is an important characteristic of ferroelectrics and the saturation polarization in the opposite direction is obtained by applying sufficiently strong field in the reverse direction reversing the direction of the domains.

7.15.2 Ferroelectric Materials

Ferroelectric materials can broadly be classified into three major groups: The Rochelle salt group, the KDP (potassium dihydrogen phosphate) group and the perovskites group headed by barium titanate (BaTiO_3). In Table 7.1, some ferroelectric materials, their curie temperature (T_c) and spontaneous polarization (P_s) are given. We briefly discuss below these three categories.

Table 7.1 Some ferroelectric materials and their properties

Material	Chemical formula	T_c (in K)	P_s (in $\mu\text{C}/\text{cm}^2$)	at $T(\text{K})$
Rochelle salt	$\text{NaK}(\text{C}_4\text{H}_4\text{O}_6)\cdot 4\text{H}_2\text{O}$	297 (upper) 255(lower)	0.026	278
KDP group	KH_2PO_4	123	4.75	96
	KD_2PO_4	213	4.83	180
	RbH_2PO_4	147	5.6	90
	CsH_2AsO_4	143	—	—

Contd.

Contd.

Perovskites	BaTiO ₃	393	26.0	296
	KNbO ₃	708	30.0	523
	PbTiO ₃	765	> 50	296
	SrTiO ₃	32	3	4.2
	LiNbO ₃	1480	71	296
	WO ₃	223	—	—

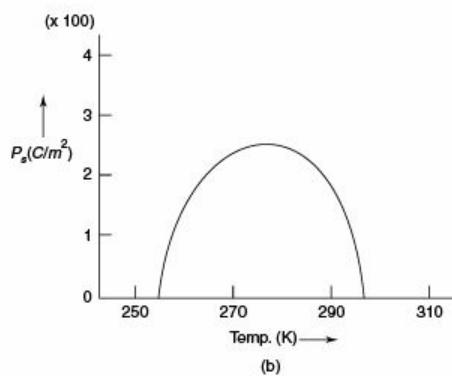
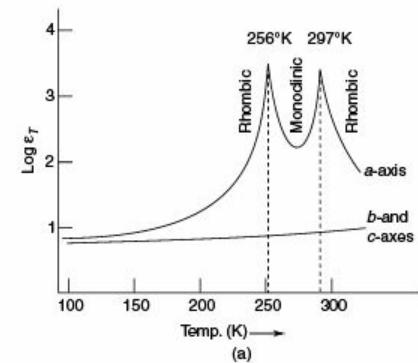


Fig. 7.16 (a) Variation of dielectric constant with temp.
 (b) Variation of P_s with temp in Rochelle shift.

7.15.3 Rochelle Salt

Rochelle salt ($\text{NaKC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$) was the first compound in which ferroelectric property was found. It was first prepared in 1672 by a pharmacist Seignatte who lived in Rochelle. The material has a unique property that it is ferroelectric only in the temperature range of 255 K to 296 K.

(-18°C to 23°C). Above 296 K and below 255 K the structure is orthorhombic while the structure is monoclinic in its ferroelectric phase. Initially, this phenomenon was known as *Seignatte electricity*. However, what we call ferroelectricity was discovered in 1920 by Joseph Valasek.

Figure 7.16 (a) and (b) respectively shows the variation of the dielectric constant and spontaneous polarization with temperature the dielectric constant peaks (~ 4000) at both the transition temperatures.

Dihydrogen Phosphate Group

Potassium dihydrogen phosphate (KH_2PO_4) is the most representative compound of this group. These are characterized by the presence of hydrogen bonds in which the motion of the protons is related to the ferroelectric properties. The curie temperature of this compound is 123 K. The variation of its spontaneous polarization is shown in Fig. 7.17.

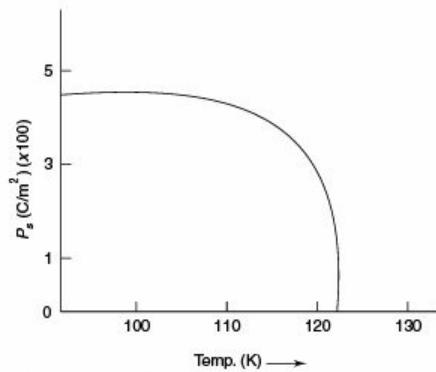
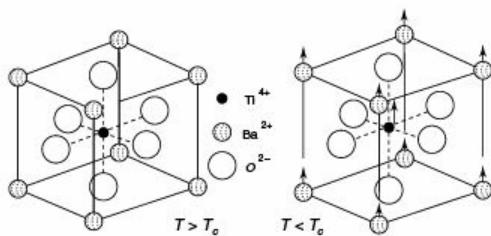


Fig. 7.17 Variation of spontaneous polarization with temperature.

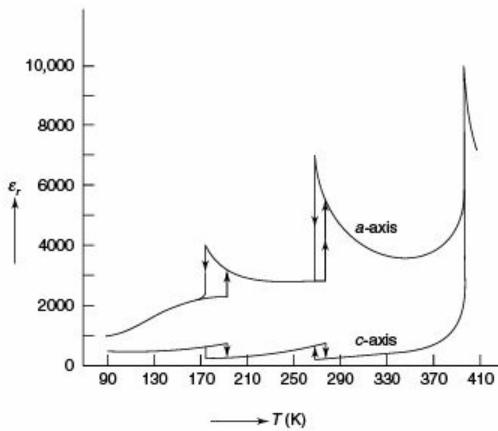
7.15.4 Perovskites Group: BaTiO_3

BaTiO_3 is the most widely investigated and used compound and is representative of perovskite type (ABO_3 structure) ferroelectric. The structure of BaTiO_3 is shown in Fig. 7.18. In the non-ferroelectric static (above T_c), it has cubic symmetry. The Ba^{2+} ions are located at the corners, the O^{2-} ions at the centre of faces and the Ti^{4+} ion is located at the centre of the cube. It has an arrangement of highly polarizable oxygen ions in the form of an octahedron with a small titanium ion at the centre. Below T_c , spontaneous polarization occurs and the compound is ferroelectric. The crystal structure becomes slightly elongated and it becomes tetragonal. The dipole moment results from the relative displacements of the Ba^{2+} , O^{2-} and Ti^{4+} ions from their symmetrical positions. Thus, a permanent ionic dipole moment is associated with each unit cell. However, at the temperature above curie temperature T_c , the unit cell becomes cubic and all ions take symmetric positions within the cubic unit cell, the unit cell now has true perovskites (ABO_3) structure and the ferroelectricity disappears.

Fig. 7.18 Crystal structures of BaTiO_3 .

BaTiO_3 has a rhombohedral structure below 193 K, orthorhombic structure between 193°–278 K and tetragonal between 278–393 K. Above 393 K it is cubic and non-ferroelectric.

In Figs. 7.19 and 7.20, the variation of dielectric constant and spontaneous polarization with temperature are shown. It is seen in Fig. 7.19 that it has three transitions at 193 K (-80°C), 278 K (5°C) and 393 K (120°C). The direction of spontaneous polarization may lie along any of the edge, i.e., there are 6 possible directions. Along the direction of spontaneous polarization of a given domain, the material expands, whereas perpendicular to polarization direction it contracts. Thus, in ferroelectric state, the material is not cubic but tetragonal. The value of P_s shown in Fig. 7.20 was measured along a cube edge (001) over the entire temperature range. Therefore, to obtain P_s in the temperature range 193–278 K the value shown in Fig. 7.31 needs to be multiplied by $\sqrt{2}$ (along (001) face diagonal). Similarly, it needs to be multiplied by $\sqrt{3}$ below 193 K (along (111) body diagonal).

Fig. 7.19 Variation of the dielectric constant of BaTiO_3 with temperature.

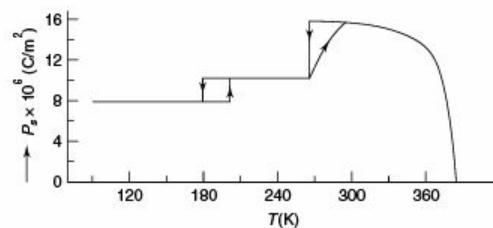


Fig. 7.20 Variation of spontaneous polarization of BaTiO_3 with temp. measured along the [001] direction.

Example 7.10

The displacement of ions in BaTiO_3 at room temperature is shown in Fig. 7.21. If the lattice constants are $a = 3.992 \text{ \AA}$ and $C = 4.036 \text{ \AA}$, calculate the magnitude of the spontaneous polarization.

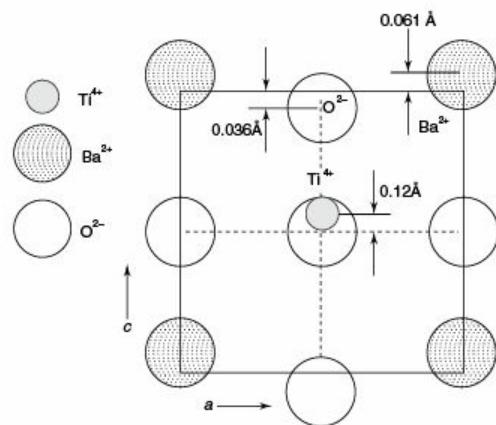


Fig. 7.21 Ionic displacement in BaTiO_3 .

Solution.

Dipole moment $p = \text{charge} \times \text{displacement}$

$$\text{Dipole moment due to the displacement of } \text{Ba}^{2+} \text{ ions} = 8 \times \left(\frac{2e}{8} \right) 0.061 \times 10^{-10} \text{ cm} \text{ (since there are 8 Ba atoms in a unit cell and each atom is shared by 8 unit cells)}$$

$= 2 \times 1.6 \times 10^{-19} \times 0.061 \times 10^{-10} \text{ cm} = 0.195 \times 10^{-29} \text{ cm}$

$$\begin{aligned} \text{Dipole moment due to the displacement of } \text{Ti}^{4+} \text{ ions} &= 4 \times 1.6 \times 10^{-19} \times 0.12 \times 10^{-10} \text{ cm} \\ &= 0.768 \times 10^{-29} \text{ cm} \end{aligned}$$

$$\begin{aligned}
 \text{Dipole moment due to the displacement of } \text{O}^{2-} \text{ ions} &= \left(-\frac{2e}{2} \right) \times (-0.036 \times 10^{-10}) \text{ cm} \\
 &= 2 \times 1.6 \times 10^{-19} \times 0.036 \times 10^{-10} \text{ cm} = 0.115 \times 10^{-29} \text{ cm} \\
 \therefore \text{Dipole moment } p &= (0.195 + 0.768 + 0.115) \times 10^{-29} \text{ cm} \\
 &= 1.078 \times 10^{-29} \text{ cm} \\
 \text{Volume of the unit cell } V &= a^2c = (3.992 \times 10^{-10})^2 \times 4.036 \times 10^{-10} \text{ m}^3 \\
 &= 64.3 \times 10^{-30} \text{ m}^3 \\
 \therefore \text{Spontaneous polarization } P_s &= \frac{p}{V} = \frac{1.078 \times 10^{-29} \text{ cm}}{64.3 \times 10^{-30} \text{ m}^3} \\
 &= 0.167 \text{ cm}^{-2}
 \end{aligned}$$

7.15.5 Antiferroelectricity

Certain dielectric crystals possess ordered electric dipoles below a characteristic temperature (called antiferroelectric curie temperature), however, do not show any spontaneous bulk polarization. In these crystal neighbouring dipoles get aligned in antiparallel direction because of which the bulk polarization of the crystal vanishes. Such crystals are called *antiferroelectric crystals* (e.g. PbZrO_3 , NaNbO_3 , Ammonium Dihydrogen Phosphate ($\text{NH}_4\text{H}_2\text{PO}_4$), etc.) and the phenomenon is termed *antiferroelectricity*.

In many perovskite type crystals, the coupling between the oxygen octahedra causes the adjacent cells to be polarized in the opposite directions. Below a certain temperature the resultant deformation is such that the total energy in the antiparallel arrangement is lower than the state of the parallel arrangement and the state of no induced dipoles. Lead zirconate (PbZrO_3) is an example. It shows two antiferroelectric phases, one ferroelectric phase and one paraelectric phase over the different temperature ranges.

7.16 ELECTRETS

Electrets are the electrical counterpart of bar magnets. A bar magnet produces a magnetic field in the space around it similarly, an electret produces an electric field in the space around it. The term electret was used for the first time by O. Heaviside in 1896. However, the first electret was obtained experimentally in 1921.

Electrets are the bar of dipolar solids, which preserve their polarization even after the removal of the polarizing electric field. These are usually prepared by polarizing the solid at high temperature in a strong electric field and then cooling it in the field and finally removing the field. This process leaves the polarization *frozen in*. Sometimes the material is melted and is allowed to solidify in the presence of a strong electric field waxes, plastics, ceramics and other inorganic or organic substances are used for making electrets.

The electrets can be used for producing electric field without an external source of voltage in, for example cathode ray tube (CRT), electrostatic generators, electrometers, etc. However, the drawback is that the electrets tend to lose their strength due to the absorption of ions from the surroundings which limits their uses.

7.17 CERAMICS

The term ceramic has its origin in the Greek word *keramikos*, which means *burnt stuff* indicating that these materials are obtained by high temperature treatment.

Ceramics are generally inorganic and nonmetallic elements such as oxides, nitrides, carbides, etc. and their production involves firing the constituents at high temperatures. The ceramics used as dielectrics are porcelains, alumina, ceramics, titanates, etc. Traditionally, the term ceramics is used for the products which use clay as raw material such as china, porcelain, bricks, tiles, etc. However, in science and technology, the term ceramic has a much broader meaning and the ceramics are used in electronics, computers, communication devices, aerospace and host of other industrial products.

Typically, ceramics are insulating to heat and electricity, brittle, hard and more resistance to high temperatures and other harsh environmental factors than the metals and polymers. Most of the technical applications of ceramics use its insulating nature towards electricity and heat. Very often ceramics are classified on the basis of the nature of application into the following groups: glasses, structural clay products, wares, refractories, cements, abrasives, electroceramics, etc. In addition to traditional uses, the development of new and what is termed *advanced ceramics* has begun and is expanding increasingly. The advanced ceramics are being used in microelectromechanical systems (MEMS), optical fibre communication systems, piezoelectric devices, high-temperature superconductors, ceramic ball bearings, etc.

7.18 ELECTROSTRICKTION

Solids, especially ceramics are relatively hard mechanically, but still expand or contract when temperature is changed or electric field is applied. The strain $\left(\frac{\Delta l \text{ (change in length)}}{l \text{ (original length)}} \right)$ caused by temperature change or stress are known as *thermal expansion* and *elastic deformation* respectively. In insulating materials, the application of an electric field also causes deformation and the resulting strain is called *electric field induced strain*.

The deformation produced in a crystal due to the application of an electric field or *electric field induced strain* is referred to as *electrostriction*. Sometimes it is mistaken as *converse piezoelectric effect* but there is difference between electrostriction and converse piezoelectric effect. Electrostriction is very often observed in ionic crystal which need not necessarily be piezoelectric. The deformation of a piezoelectric crystal is proportional to the applied electric field and the stress induced polarization varies linearly with the strain produced. The converse piezoelectric effect is a primary electromechanical coupling effect, i.e., the strain is proportional to the applied electric field, while the *electrostriction is a secondary coupling in which the strain is proportional to the square of the electric field*.

The converse piezoelectric effect is expressed as

$$x = dE \quad \dots(7.63)$$

where x is the strain (a unit cell length change), E the electric field and the proportionality constant d is called the piezoelectric constant. The electrostrictive effect is expressed as

$$x = ME^2 \quad \dots(7.64)$$

where M is the electrostrictive constant. The electrostrictive materials do not have a spontaneous polarization and so do not generate any charge under stress but exhibit a change in permittivity. The *converse electrostrictive effect*, i.e., the stress dependence of the permittivity is used in *stress sensors*. Unilike piezoelectric sensors, electrostrictive sensors are effective in the low frequency range particularly d.c. fields.

7.19 DIELECTRIC STRENGTH/ BREAKDOWN

A dielectric possesses a very small electrical conductivity and at ordinary applied electric field do not conduct electricity. However, if the magnitude of the applied electric field is increased beyond a certain critical value, the insulating property of the dielectric is lost and the conductivity increases suddenly. This phenomenon is called *dielectric breakdown* and the corresponding electric field (or voltage) is referred to as *dielectric (break down) strength*. The dielectric strength is also defined as the *maximum potential gradient to which a dielectric can be subjected with insulation breakdown*. That is,

$$\text{Dielectric strength} = \left(\frac{dV}{dx} \right) = \frac{V_B}{d} \quad \dots(7.65)$$

where V_B is the breakdown voltage and d is the thickness of the specimen. In Table 7.2, few dielectric materials and their dielectric strength are given. Mica has one of the highest dielectric strengths and is widely used for insulation purposes.

Table 7.2

Dielectric material	Dielectric strength
Mica	100×10^6 V/m
Bakelite	15×10^6 V/m
Transformer oil	10×10^6 V/m
Silica	10×10^6 V/m
Stealite	12×10^6 V/m
Porcelain	6×10^6 V/m

The factors which determine the dielectric strength of a material are its composition, microstructural features (porosity, cracks, flaws), thickness, temperature and humidity. Dielectric breakdown occurs in gases due to corona discharge in non-uniform fields or due to photoionization and collision of electrons with gas atoms, leading to further ionization, increased conduction and subsequently dielectric breakdown. In liquids, breakdown is primarily due to dielectric heating and increased ionization processes. The process is greatly accelerated by the presence of impurities or high concentration of mobile ions (e.g. Na^+ , Li^+) in the liquid. For solids, the main breakdown mechanisms are intrinsic, thermal and defect breakdowns. Another mechanism, called electrochemical degradation or breakdown, is associated with ion migration under a (d.c.) field gradient at higher temperatures.

7.19.1 Intrinsic Breakdown

At very high applied electric fields, the electrons may be ejected from the electrode materials or due to the excitation of electrons from valence band to conduction band. These electrons are then

accelerated through the material by the high electric field which undergo further collisions producing more electrons resulting in an 'avalanche' of conducting electrons over a narrow area and the dielectric breakdown occurs. Such a breakdown is called *intrinsic* or *electronic breakdown* and can occur in all solids at extremely large fields ($\sim 10^8$ V/m). The presence of impurities can create additional energy levels in the energy gap and help in excitation of electrons into conduction band. An increase in temperature aids the thermal excitation increasing the possibility of the intrinsic breakdown.

7.19.2 Thermal Breakdown

Thermal breakdown occurs due to local heating or conduction losses which generate heat at the rate faster than it can be dissipated. This increases the local temperature even upto melting point or evaporation. Depending on the material and the thermal gradients formed, cracks develops around the hot spot. The combination of ionized gases and increased conduction cause dielectric breakdown. This is the most common dielectric failure mode in ceramic insulators.

7.19.3 Defect Breakdown

In inhomogeneous dielectric solids having structural defects such as cracks and pores, the ionization of gases inside the pores takes place when subjected to high electric field. This results in the local generation of heat causing a temperature gradient. This leads to thermal stresses which in turn generate cracks leading to further ionization and subsequent breakdown. This breakdown is called *defect breakdown* or *ionization breakdown*. The presence of porosity in dielectrics is therefore detrimental for higher dielectric strengths.

7.19.4 Electromechanical Breakdown

Electrochemical breakdown is usually caused by ion (e.g. Na^+ , Ag^+ , OH^-) migration under a continuously high applied d.c. field (30 kV/m) at moderately elevated temperatures (50 to 250°C). This breakdown can cause interelectrode shorting of closely spaced electrodes or depletion at the one or both the electrodes or ion build up leading to high local field and electrode corrosion. This can occur in microelectronic circuits such as Ag conducting wires on an Al_2O_3 substrate when a potential difference exist between the lines. The presence of high humidity can aggravate the situation and so encapsulation of such circuits is often necessary.

To avoid breakdown dielectrics are often *glazed* to make the surface non-absorbent and free cracks and pores. The presence of cracks and pores can lead to ionization of entrapped gases at high fields particularly under high humidity conditions, resulting in the dielectric breakdown. Also, to decrease the possibility of surface shorting, insulators are designed with lengthened surface paths.

SHORT ANSWER TYPE QUESTIONS

1. What is a dielectric? What is an ideal dielectric?
2. Is there any difference between an insulator and a dielectric? Explain.
3. Define polarization. Briefly explain its mechanism.
4. Differentiate between polar and non-polar dielectrics.
5. Define 'dielectric constant'? Explain what this parameter indicates about a material.

6. Define electric susceptibility and polarizability. What these parameters are indicative of?
7. Show $P = \epsilon_0 (\epsilon_r - 1) E$, where P is the electric polarization.
8. What do you mean by local electric field? Discuss.
9. Why orientational polarization is also known as dipolar polarization? Is it temperature dependent or independent? Explain.
10. Explain ionic polarization. Why this is also known as 'displacement' polarization.
11. Explain why optical dielectric constant is significantly smaller than the static dielectric constant?
12. What do you mean by dielectric loss? Explain the meaning of $\tan \delta$.
13. Name the factors which determine the dielectric loss.
14. What do you mean by 'smart materials'? Name a few small dielectric materials.
15. Define piezoelectric effect. What is the necessary condition for the existence piezoelectricity in a crystal? Give few examples of these materials.
16. Define 'spontaneous polarization'.
17. What do you mean by pyroelectricity? How pyroelectric materials are useful?
18. What are ferroelectric materials? Name a few ferroelectric materials. What are their main characteristics?
19. Define ferroelectric curie temperature. What is its value of BaTiO_3 ? When does a ferroelectric materials turn into a paraelectric material?
20. Discuss the applications of ferroelectric materials.
21. Explain antiferroelectricity. Name a few antiferroelectric materials.
22. What are electrets? Are they different from ferroelectric materials?
23. What are "ceramics"? How they are useful to us?
24. Explain electrostriction.
25. Define dielectric breakdown. Differentiate between intrinsic breakdown and thermal breakdown.
26. What precautions are taken to avoid dielectric breakdown? Explain.
27. Show that the surface charge density of the polarization charges on the outer surface of a dielectric is given by

$$\sigma = P \times \hat{n}$$

where \hat{n} is a unit vector normal to the surface.

LONG ANSWER TYPE QUESTIONS

1. Explain the polarization phenomenon in dielectrics. Obtain a relation between dielectric constant and polarizability.
2. Obtain Claussius-Mossotti relation and mention its significance.
3. Name various types of polarization and explain their mechanism. Show their frequency dependence and discuss.
4. What is electronic polarization? Show that in a static electric field $a_e = 4\pi \epsilon_0 r^3$, where r is the radius of the atom.
5. What is meant by dielectric loss? Obtain an expression for the same.
6. Discuss piezoelectric effect and the converse piezoelectric effect. Explain the condition for existence and the applications of this effect.

7. Discuss the characteristics of ferroelectric materials. Draw the polarization versus electric field curve for a ferroelectric material and explain remanent polarization and coercive electric field.
8. Explain the dipolar theory of ferroelectricity.
9. What is electrostriction? Differentiate between electrostriction and piezoelectric effect.
10. What do you mean by dielectric breakdown? Explain various types of dielectric breakdowns.

NUMERICALS

1. The relative dielectric constant of polystyrene is 2.5 when subjected to a d.c. field. Find the polarization when a 0.5 mm thick polystyrene sheet separates 100 volt.
[Ans. $2.7 \times 10^{-6} \text{ cm}^{-2}$]
2. The relative dielectric constant of a glass and of a plastic are 3.9 and 2.1 respectively. What voltage should be applied between the electrodes separated by 0.13 cm of glass, if we want the same charge density as would develop on another set of electrodes separated by 0.42 cm of plastic and 210 volts.
[Ans. 3.5 volt]
3. A parallel plate capacitor of plates $1 \text{ cm} \times 2 \text{ cm}$ each, receives a potential difference of 2.25 volts between the electrodes. (a) How far apart must the plates be to produce a charge density of 10^{-7} cm^{-2} ? (b) How many electrons accumulate on the negative plate?
[Ans. (a) 0.2mm (b) 1.25×10^8]
4. The dielectric constant of Al_2O_3 is 8. Find the capacitance of a layer of Al_2O_3 of thickness 0.5 μm and surface area 2000 mm^2 .
[Ans. $2.8 \times 10^{-7} \text{ F}$]
5. Calculate the shift of the electronic cloud with respect to the nucleus in an argon atom when subjected to an electric field of 10^5 V m^{-1} .
[Ans. $5 \times 10^{-18} \text{ m}$]
6. If the polarizability of helium is $0.18 \times 10^{-40} \text{ Fm}^2$, find its dielectric constant at NTP.
[Ans. 1.00055]
7. If the polarizability of a Kr atom is $2.18 \times 10^{-40} \text{ Fm}^2$, calculate the dielectric constant of Kr at 0°C and at one atmosphere.
[Ans. 1.00066]
8. The susceptibility of CO_2 is 0.985×10^{-3} and its density is 1.977 kg m^{-3} . Find its polarizability.
[Ans. $3.24 \times 10^{-40} \text{ Fm}^2$]
9. The atomic polarizability of helium is $0.18 \times 10^{-40} \text{ Fm}^2$. Calculate the induced dipole moment per unit volume (density of polarization) of He gas if it is subjected to an electric field of strength $6 \times 10^6 \text{ V m}^{-1}$. The concentration of helium atoms is $2.6 \times 10^{25} \text{ m}^{-3}$. Also find the separation between the positive and negative charges in each atom.
[Ans. $2.81 \times 10^{-10} \text{ cm}^{-2}, 3.4 \times 10^{-17} \text{ m}$]
10. The polarizability of neon gas is $0.35 \times 10^{-40} \text{ Fm}^2$. If the gas contains 2.7×10^{25} atoms per m^3 at 0°C and one atmosphere, calculate its dielectric constant.
[Ans. 1.000108]
11. The following data is given about a C Cl_4 molecule
Relative permittivity, $\epsilon_r = 2.24$

Density $P = 1.6 \times 10^3 \text{ kg m}^{-3}$

Molecular weight $M = 156$

If it is subjected to a field of strength 10^7 Vm^{-1} , Calculate the dipole moment of a molecule.
Also find the average electron displacement. [Ans. $1.77 \times 10^{-32} \text{ cm}$, $1.5 \times 10^{-15} \text{ m}$]

Hint: Molecular density $N = \frac{\text{Avogadro's Number} \times \text{Density}}{\text{Molecular weight}}$

Dipole moment $\mu_o = \frac{P}{N} = \frac{\epsilon_o (\epsilon_r - 1)E}{N}$

Displacement $x = \frac{\mu_o}{Ze}$

12. A sample of NaCl containing $1.6 \times 10^{20} \text{ molecules/m}^3$ is subjected to an electric field of strength $5 \times 10^5 \text{ Vm}^{-1}$. Calculate the orientational polarization at room temperature.

Take the separation between Na^+ and Cl^- ions to be 0.25 nm. [Ans. 10^{-11} cm^{-2}]

8

Superconductivity

8.1 DISCOVERY OF SUPERCONDUCTIVITY

The current in a conductor is carried by conduction (or free) electrons. According to the concepts of quantum mechanics, a moving electron can be represented by a plane wave progressing in the direction of motion of the electron. A metal has a crystalline structure with atoms occupying periodic lattice points. A plane wave has the property that it can pass through a perfect periodic structure without any scattering (or collision) into other directions. Hence, an electron is able to pass through a perfect crystal without any loss of momentum in its direction of motion. In other words, *the electrical resistance in a perfect crystalline defect free material is zero*. However, for the following reasons, every material has always a finite resistance. No material is hundred per cent pure and the presence of defects or impurities interrupt periodicity. Moreover, at any temperature, other than absolute zero, the atoms of a material vibrate and are displaced by different amounts from its equilibrium position, depending upon the temperature of the material. These vibrations and impurities (or imperfections or defects) cause the scattering of moving conduction electrons and give rise to electrical resistance.

When the temperature of a metal or alloy is lowered, the magnitude of thermal vibrations of the atoms decrease causing lesser scattering of conduction electrons resulting in the decrease of electrical resistance. The decrease of resistance is linear down to a temperature to nearly one-third of the characteristic Debye temperature of the material and below this temperature the decrease in resistance is less rapid (Fig. 8.1). For a hypothetical pure metal, the scattering of the conduction electrons is only due to the thermal vibrations of the lattice and so the resistance is expected to drop to zero as the temperature is reduced towards 0 K. However, this phenomenon is not superconductivity. In fact, no material is hundred per cent pure and always has some amount of defects and impurities which cause scattering of conduction electrons even at the lowest possible temperature. Therefore, even at the lowest temperature, every conductor has certain resistance or resistivity called *residual resistivity*, ρ_o . The more impure a conductor, higher is its residual resistivity.

Kammerling Onnes Experiment

The liquification of helium in 1908 made it possible to investigate the physical properties of the

material in the vicinity of absolute zero. At such low temperatures, the thermal vibrations of atoms are almost completely ceased. K. Onnes, a German scientist, at Leiden in 1911, decided to investigate the behaviour of electrical resistance of a metal at extremely low temperature using liquid helium as coolant. He took mercury as the metal for his experiment because it was available in the purest form, as it could be purified by repeated distillation. He observed that as the mercury was cooled from 233 K to 4.2 K resistance decreased almost linearly by a factor of 500.

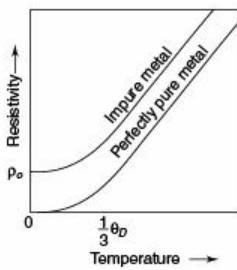


Fig. 8.1 Variation of resistance of metals with temperature.

However, at 4.2 K, in the words of Onnes, “*within some hundredths of a degree came a sudden fall, not foreseen by the vibrator theory of resistance, bringing the resistance at once less than a millionth of its original value ...*” Mercury has passed into a new state (Fig. 8.2), which on account of its extraordinary electrical properties may be called ‘superconducting state.’ Thus, the term superconductivity was coined by K. Onnes who discovered this wonderful phenomenon in which the electrical resistance of a material drops to immeasurably small value. Such materials are called *superconductors* and the temperature at which the material becomes superconductor is called its critical transition temperature (T_c). K. Onnes was awarded Nobel Prize in Physics in 1913 for his discovery of superconductivity.

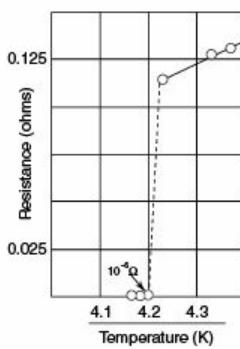


Fig. 8.2 Variation of resistance in Hg as observed by Onnes leading to the discovery of superconductivity.

Onnes and co-workers found superconductivity in many metals, e.g., niobium ($T_c = 9.3$ K), lead ($T_c = 7.2$ K), tin ($T_c = 3.7$ K), etc. Niobium is the metallic element with the highest T_c . More than twenty metallic elements and large number of compounds have been found to exhibit superconductivity. Table 8.1 gives some of the elements and compounds showing superconductivity and their transition temperatures.

Table 8.1 Some superconducting elements and compounds and their T_c

Material	T_c (K)
Hg	4.2
Nb	9.3
Pb	7.2
In	3.4
Sn	3.7
Al	1.2
V	5.4
Nb_3Ge	23.2
Nb N	16.0
Nb_3Sn	18.5
Nb_3Al	18.0
V_3Ga	16.5
$\text{K}_3\text{C}_6\text{O}$	19.2
V_3Si	17.1
$\text{YBa}_2\text{Cu}_3\text{O}_7$	91
$\text{Bi}_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$	110
$\text{Tl}_2\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$	125

Figure 8.3 shows the superconducting transition in pure and impure form of tin. Whether the transition from normal to superconducting state is sharp or not depends on the purity of the material. When the specimen is pure the transition is very sharp. However, when the same material is not so pure or has defects in crystal structure the transition from normal to superconducting state broadens and the transition is not sharp.

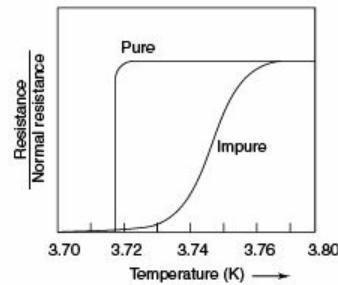


Fig. 8.3 Superconducting transition in tin.

Since the discovery of superconductivity in Hg at 4.2 K, scientists have been trying tirelessly to discover superconductors at room temperature, for its fascinating impact on technology—discussed later in this book. Almost after a century, this journey towards room temperature superconductors has reached to 181 K, the transition temperature of $\text{Sn Pb}_{0.4} \text{In}_{0.6} \text{Ba}_4 \text{Tm}_5 \text{Cu}_2 \text{O}_{20}$. Let us hope a room temperature superconductor is discovered in near future.

8.2 ZERO RESISTANCE - PERSISTENT CURRENT

Often a superconductor is defined as a material having zero resistance. The important question is how one can be sure that the resistance of a material has fallen to zero when it undergoes superconducting transition. One can only say for sure that the resistance has fallen below the smallest measurable value (or sensitivity) of the instrument employed. However, no experiment has been able to detect any resistance in the superconducting state of a material.

If a metallic conductor (even of silver, the best metallic conductor) is connected to a source of emf, after some time, depending upon the capacity of the emf source, the current in the loop reduces to zero, due to dissipation of power owing to joule heating. However, in a loop of superconducting wire, there is no decay of current for the years, i.e., *persistent electrical current flow without any attenuation*. J. File and R.G. Mills have studied the decay of current in a superconducting loop using precision nuclear magnetic resonance (NMR) method.

If L is the self-inductance of the loop and at time $t = 0$ the current is $i(0)$ flowing in the loop, then the current at a later time t decays according to the relation

$$i(t) = i(0) e^{-(R/L)t} \quad \dots(8.1)$$

where R is the resistance of the loop. They measured the magnetic field produced by the circulating current. The measurement of the magnetic field does not draw energy from the circuit and it is possible to find out whether the current circulates indefinitely. Gallop found that there is hardly any decay of current circulating round a closed loop of superconducting wire. He also estimated the resistivity of the superconducting wire to be 10^{-26} ohm-m, which is less than 10^{-18} times than that of copper at room temperature. These observations justify the assumption that the resistance of a superconductor is zero.

8.3 A.C. RESISTIVITY IN SUPERCONDUCTORS

The absence of any resistance across a superconducting specimen means that there is no voltage drop along the metal when a current is passed through it and so no power is generated by the passage of the current. This, however, is strictly true only for a direct current having a constant magnitude. If the current is changing an electric field is developed leading to the dissipation of power. In order to understand this, we discuss below the behaviour of conduction electrons in the superconductors.

Below the superconducting transition temperature, there are two types of electrons in a superconducting specimen. Those electrons which take part in superconductivity (exist as pairs and known as cooper pairs), also called *super electrons*, can pass through the material without resistance because they do not undergo scattering or collision. The remaining electrons, behaving as normal electrons, undergo scattering and so encounter resistance as in a normal material. As the temperature is lowered below T_c , the fraction of superelectrons increases. At 0°K, all the electrons behave as superelectrons. On the other hand, as the temperature is raised towards T_c , more and more electrons

start behaving as normal electrons and just above T_c all the electrons become normal electrons and the material becomes non-superconducting. Thus, a superconductor below its transition temperature appears to be permeated by two electron fluids, one of normal electrons and the other of superelectrons. The relative electrons density in the two fluids depends upon the temperature. The *two-fluid model* is also supported by thermodynamic considerations.

In a superconductor, in general, current is carried by both superconducting and normal electrons. However, when a direct current of constant magnitude is flowing, the current is entirely due to superelectrons. We can understand this fact as if the current is to remain constant, there must be no electric field in the material otherwise the superelectrons would be accelerated continuously and the current would increase indefinitely. If there is no electric field, there is nothing to drive the normal electrons and there is no normal current. Therefore, for a constant value of current, all the current is carried by the superelectrons.

A superconducting specimen can be considered to be equivalent to two conductors in parallel, one having zero resistance and the other having normal resistance. The superelectrons taking the path of zero resistance conductor do not encounter any resistance while the normal electrons passing through the normal conductor encounter finite resistance. We can also visualize the situation in the following manner. If we suddenly apply voltage across a superconducting specimen, the current tends to rise to infinity but it is limited by the internal resistance of the source. While the current is changing, an electric field must be present to accelerate the electrons. Electrons, however, possess certain inertial mass and so the supercurrent does not rise instantaneously. If an accelerating voltage is applied, the supercurrent lags behind the field because of the inertia of the superelectrons. Thus, *the superelectrons present an inductive impedance*. However, due to the presence of an accelerating electric field, some of the current is also carried by the normal electrons. Thus, current is not carried entirely by superelectrons as in the case of a d.c. flow. Although, normal electrons also possess inertial mass and so some inductive reactance, however, the inductive component is completely overshadowed by resistance resulting from scattering (or normal resistance). Thus, the resistive behaviour of a superconducting specimen can be explained by a perfect inductance in parallel with a resistance.

The normal component of the current dissipates power in the usual way. However, due to very small inertial mass of electron, the inductance due to their inertia is extremely small and the resulting power dissipation is also small. Nevertheless, with a.c. signal, a superconducting specimen is not resistance free as in the case of a d.c. signal.

At extremely high frequencies, however, a superconductor responds in the same way as a non-superconducting material. This can be understood by the following discussion. The superelectrons are in lower energy state than the normal electrons but if the frequency of the applied field is high enough, the superelectrons can get excited into higher energy state where they behave as normal electrons. This happens for frequencies greater than about 10^{11} Hz, higher than the frequencies of very long wave infrared. The behaviour of a superconductor at optical frequencies is, therefore, no different from that of a normal material. For example, there is no change in the visual appearance of a superconductor as it is cooled below its transition temperature.

8.4 EFFECT OF MAGNETIC FIELD: CRITICAL MAGNETIC FIELD

On the application of a sufficiently strong magnetic field to a superconducting specimen, superconductivity disappears from the specimen and it becomes a normal or non-superconducting

specimen. The minimum applied magnetic field at which a superconducting specimen loses its superconductivity is called the critical magnetic field of the specimen, and is denoted by $H_c(T)$. As the temperature of the specimen approaches T_c from lower temperature, the value of H_c approaches zero from a higher value. It is a function of temperature, i.e., its value depends upon the temperature of the superconducting specimen. The experimental variations of critical magnetic field of some of the elements are shown in Fig. 8.4. The curves separate the superconducting region from the non-superconducting region. The curves are closely approximated by parabola of the form shown in Fig. 8.5 and can be mathematically expressed as:

$$H_c(T) = H_{co} \left(1 - \frac{T^2}{T_c^2} \right) \quad \dots(8.2)$$

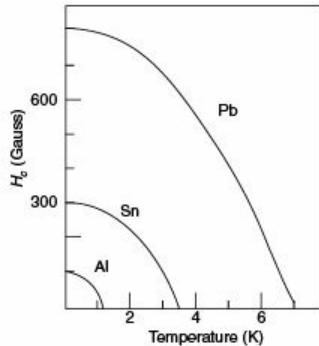


Fig. 8.4 Variation of H_c with temperature for some elemental superconduction.

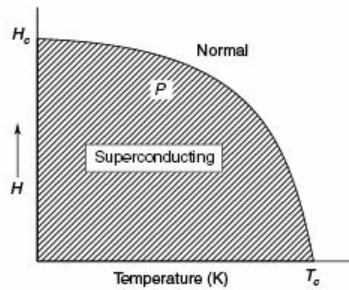


Fig. 8.5 Magnetic phase diagram of a superconductor.

where H_{co} is the critical magnetic field 0 K and $H_c(T)$ the critical magnetic field at T K. Although, the experimental curves are not exactly parabolic, for the sake of generalization, the general behaviour can be approximated by a parabola. Each superconductor can be characterized by H_{co} and T_c . Knowing

H_{∞} and T_c , the value of $H_c(T)$ can be estimated by using Eq. (8.2). Table 8.2 mentions the value of H_{∞} for some of the elements.

Table 8.2 Critical magnetic field H_{co} of some of the elements

Element	H_{co}	
	(A/m)	Gauss
Al	7.9×10^3	99
Cd	2.4×10^3	30
Ga	4.1×10^3	51
Pb	64×10^3	803
Hg (α)	33×10^3	413
Hg (β)	27×10^3	340
Zn	3.8×10^3	53
V	105×10^3	1020
Ta	66×10^3	830
Sn	24×10^3	306

The diagram, shown in Fig. 8.5, is often called the phase diagram of a superconductor. If the value of H_c and T is such that it is below the curve, e.g., a point P , lying in the shaded region, then the material is in superconducting phase. If the value of H_c and T is such that it corresponds to a point outside the shaded region, the material is in non-superconducting or normal state. The arrows in the figure indicate that the specimen can be driven from superconducting to normal state either by increasing the temperature or increasing the applied magnetic field or both.

K. Onnes and his co-workers had found that when an electric current through a superconducting wire exceeds a specific value the superconductivity is lost. This specific value of current is called the *critical current* (I_c) for the particular superconductor. As we know, a current carrying conductor generates magnetic field around it. According to Silsbee, the two phenomena (i.e., destruction of superconductivity by magnetic field and electric current) are related. Silsbee suggested that the important factor in the destruction of superconductivity is the magnetic field produced by the passing current rather than the current itself. Thus, *the destruction of superconductivity is field controlled*.

Critical current is related to the critical magnetic field strength. When the current strength reaches the critical current I_c , the associated magnetic field have value equal to critical magnetic field

$$H_c = 2\pi r I_c \quad \dots(8.3)$$

where, r is the radius of the circular conductor.

Example 8.1

The critical magnetic field of a superconducting specimen at 12 K is 15 T while that at 10 K is 18 T. What is the critical transition temperature of the specimen? Also, calculate the critical field at 0 K.

Solution.

Here, at

$$T = 12 \text{ K} \quad H_c = 15 \text{ T}$$

and at

$$T = 10 \text{ K} \quad H_c = 18 \text{ T}, \quad T_c = ?, H_{\infty} = ?$$

Substituting these in

$H_c(T) = H_{co} \left(1 - \frac{T^2}{T_c^2}\right)$, we get

$$H_c(12) = H_{co} \left(1 - \frac{(12)^2}{T_c^2}\right) = 15 \quad \dots(i)$$

and $H_c(10) = H_{co} \left(1 - \frac{(10)^2}{T_c^2}\right) = 18$

Dividing the expressions, we have

$$\frac{H_c(12)}{H_c(10)} = \frac{1 - \frac{144}{T_c^2}}{1 - \frac{100}{T_c^2}} = \frac{15}{18}$$

or $\frac{T_c^2 - 144}{T_c^2 - 100} = \frac{15}{18} = \frac{5}{6} \Rightarrow 6 T_c^2 - 864 = 5 T_c^2 - 5000$

or $T_c^2 = 364 \Rightarrow T_c \approx 19 \text{ K (Approx.)}$

Also, from (i) we have, $H_c(0) = \frac{15}{1 - \frac{144}{T_c^2}} = \frac{15}{1 - \frac{144}{364}} = \frac{15}{\frac{220}{364}} = 24.81 \text{ K}$

Example 8.2

The critical temperature of a superconducting is 20 K. At what temperature its critical magnetic field will reduce to one-third of its value at 0 K?

Solution.

Here, $H_c(T) = \frac{H_{co}}{3}$ at $T = ?$

$$H_c(T) = H_{co} \left(1 - \frac{T^2}{T_c^2}\right)$$

$\therefore \frac{H_{co}}{3} = H_{co} \left(1 - \frac{T^2}{T_c^2}\right)$

$$1 - \frac{T^2}{T_c^2} = \frac{1}{3} \Rightarrow \frac{T^2}{T_c^2} = 1 - \frac{1}{3} = \frac{2}{3} \Rightarrow T = \sqrt{\frac{2}{3}} \cdot T_c = \sqrt{\frac{2}{3}} \cdot 20 = 16.32 \text{ K}$$

Example 8.3

Find the critical current and critical current density of a thin superconducting wire of diameter 0.001 m if the critical magnetic field of the wire is $7.9 \times 10^3 \text{ A/m}$.

Solution.

We know, critical current $I_c = 2\pi r H_c$

Here, $2r = 0.001 \text{ m} = 10^{-3} \text{ m}$

$H_c = 7.9 \times 10^3 \text{ A/m}$

$$\therefore I_c = \frac{22}{7} \times 10^{-3} \times 7.9 \times 10^3 \text{ A} = 24.806 \text{ A}$$

$$\text{Critical current density } J_c = \frac{I_c}{\text{Area}} = \frac{I_c}{\pi r^2} = \frac{24.806}{\frac{22}{7} \times (0.0005)^2} = 3.16 \times 10^7 \text{ A/m}$$

8.5 MEISSNER EFFECT – FLUX EXCLUSION

Meissner and Oschenfeld independently found in 1933 that a superconductor completely excludes an externally applied magnetic field. This phenomenon is known as *Meissner effect*. The effect is of fundamental importance as it shows that a bulk superconductor behaves in an external magnetic field H as in the interior of a superconductor

$$B = \mu_0 (H + M) = 0$$

$$\text{or magnetic susceptibility } \chi = \frac{H}{M} = -1$$

That is, a superconductor exhibits *perfect diamagnetism*. In fact, no normal diamagnetic material is perfectly diamagnetic. Only a superconductor in its superconducting state is perfectly diamagnetic. Here, it is assumed that the strength of the applied magnetic field is less than the critical magnetic field of the superconducting specimen. Figure 8.6 shows that the behaviour of a superconducting specimen and Fig. 8.7 that of a non-superconducting specimen in a magnetic field on the cooling them (less than T_c for superconducting specimen). It can be observed that while a superconductor in its superconducting state never permits a magnetic flux inside it, it is not so for non-superconducting perfect conductor.

The perfect diamagnetism of superconductors is an independent property not at all related to zero resistivity as revealed from the following.

From Ohm's law $E = \rho J$, in a superconducting specimen as resistivity ρ goes to zero, keeping current density J finite, the electric field must also be zero. From Maxwell's equation (Faraday's law), we have

$$\frac{d\mathbf{B}}{dt} = -\operatorname{curl} \mathbf{E} \quad \dots(8.4)$$

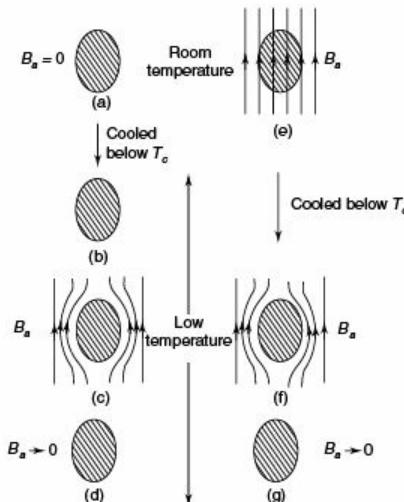


Fig. 8.6 Magnetic behaviour of a superconductor, (a) → (b) Specimen turns resistanceless in absence of magnetic field. (c) Magnetic field applied to superconducting specimen. (d) Magnetic field removed. (e) → (f) Specimen becomes superconducting in applied magnetic field (g) Applied magnetic field removed.

From the above discussion, it is clear that for a material with zero resistivity (as $E = 0$)

$$\frac{dB}{dt} = 0 \quad \text{...}(8.5)$$

or $B = \text{constant}$

In other words, the magnetic flux through the specimen should remain unchanged on cooling through the transition. Obviously, this is in contradiction with the Meissner effect. However, both Meissner effect and zero resistivity are experimentally observed facts.

The paradox was resolved by the realization of the fact that *zero resistivity and perfect diamagnetism are the two independent and essential properties of the superconducting state. However, it is the Meissner effect which is taken as the sure signature of superconductivity*, as no material other than superconductors shows total flux exclusion or perfect diamagnetism. While the resistance of a perfectly pure ideal conductor can approach zero near absolute zero, at least theoretically.

When a magnetic field is applied to a superconducting specimen, a resistanceless current is induced which circulates on the surface of the specimen in such a manner so as to create magnetic flux density everywhere inside the specimen such that it is exactly equal and opposite to the flux density of the applied magnetic field. The circulating surface currents are often termed *screening currents*.

$$\text{Flux density} \quad B = \mu_0 (H + M).$$

$$\text{For a superconductor} \quad B = 0$$

$$\therefore M = -H$$

That is, a superconducting material behaves as if it acquires a negative bulk magnetization.

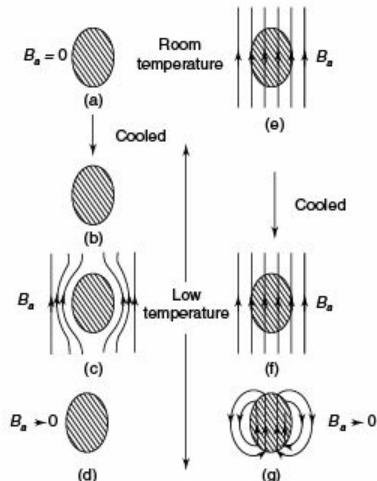


Fig. 8.7 Magnetic behaviour of a "perfect" conductor. (a) → (b) Specimen becomes resistanceless in absence of field. (c) Magnetic field applied to resistanceless specimen (d) Magnetic field removed. (e) → (f) Specimen becomes resistanceless in applied magnetic field (g) Applied magnetic field removed.

8.6 TYPE I AND II SUPERCONDUCTORS

On the basis of the behaviour of superconductors, the applied magnetic field, the superconductors are classified into two categories, i.e., *type-I superconductors* and *type-II superconductors*. The variation of magnetization of a type-I superconductor is shown in Fig. 8.8 (a) and that of type-II superconductor is shown in Fig. 8.8 (b).

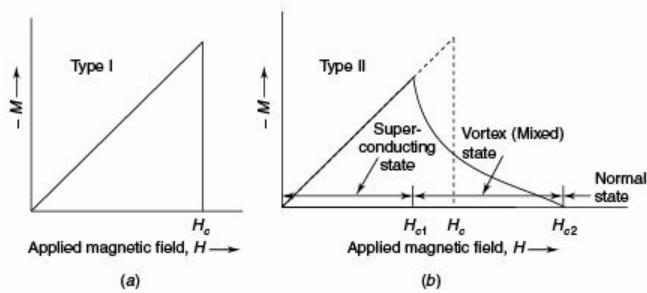


Fig. 8.8 (a) Magnetization versus magnetic field behaviour of (a) Type-I and (b) Type-II superconductors.

Type-I superconductors are characterized by a sharp critical magnetic field below which it is perfectly diamagnetic, i.e., the flux density within the specimen is zero. However, when the applied magnetic field strength reaches the critical value, H_c , the superconductor is driven into normal state and the flux within the specimen is no longer zero. For all higher magnetic fields it remains non-superconducting. For $H \geq H_c$, the negative magnetization disappears and there is no magnetization left in the material. Most of the elemental superconductors are type-I superconductors. The magnetic phase diagram of a type-I superconductor is shown in Fig. 8.5.

Type-II superconductors are characterized by two values of critical magnetic fields, the *lower critical magnetic field* H_{c1} and the *upper critical magnetic field* H_{c2} . For the applied magnetic field upto H_{c1} , the behaviour is exactly identical to that of type-I superconductors, i.e., there is total flux exclusion. At H_{c1} , the magnetic flux begins to penetrate the specimen and the penetration increases continuously until the upper critical magnetic flux H_{c2} is reached. At H_{c2} , the negative magnetization vanishes and the specimen becomes non-superconducting. In between H_{c1} and H_{c2} the specimen is in *vortex* (or *mixed*) state. The value of H_{c2} may be as high as 100 times more than H_{c1} . The *basic difference is that the magnetization of these superconductors vanishes gradually as the magnetic field is increased rather than sharply as in type-I superconductor*. However, as far as electrical conduction is concerned, a type-II specimen is completely superconducting at all magnetic fields below H_{c2} . Most of the alloys and transition metals show type-II superconductivity. The magnetic phase diagram of a type-II superconductor is as shown in Fig. 8.9.

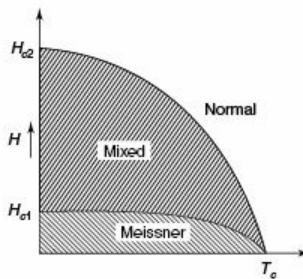


Fig. 8.9 Magnetic phase diagram of a type II superconductor.

Due to flux exclusion, the flux is concentrated at the sides of the specimen. This means that the field at the sides of the specimen is higher than the top and bottom. Therefore, the critical value H_c is reached at sides earlier than that at top and bottom. This results in the initiation of transition to normal state from superconducting state at the sides before the top and bottom. As a result, the specimen exist as a complex mixture of normal and superconducting regions called intermediate or vortex state.

The value of H_c is too low for any practical application. But for the type-II superconductor H_{c2} can be higher than 10 T. It is 54 T for PbMO_6S_8 while for cuprate (high-temperature superconductors) it can be as high as 150 T. High values of H_{c2} have made it possible to use these superconductors in Magnetic Resonance Imaging (MRI) for medical applications.

The behaviour of a superconducting specimen can change from type-I to type-II on adding some impurities. For example, Pb shows type-I behaviour ($H_c = 4.8 \times 10^4$ A/m at 4 K) but on adding 2% indium in it shows type-II behaviour with $H_{c1} = 3.18 \times 10^4$ A/m and $H_{c2} = 7.96 \times 10^4$ A/m.

Ideal and Non-ideal Superconducting Specimen: If the magnetization curve of a superconducting specimen is *reversible*, the specimen is said to be *ideal*. While a superconducting specimen whose magnetization curve is *non-reversible* is called *non-ideal superconducting* specimen. Both type-I and type-II superconductors can be ideal or *non-ideal*. Pure specimen free from crystalline defects normally show ideal behaviour. Type-II superconductors with a large amount of magnetic hysteresis induced by mechanical treatment are called *hard superconductors*.

8.7 PENETRATION DEPTH

We have seen that magnetic field lines cannot penetrate through a superconductor. Perfect diamagnetism of a superconductor prevents electric currents from flowing through the body interior of the specimen and the current is confined near the surface. However, current cannot be confined entirely to the surface otherwise the current density would rise to infinity, which is physically impossible. The current in a superconductor flows within a very thin surface layer whose thickness is of the order of 10^{-5} cm, the exact value depending upon the material and its temperature.

When a superconducting specimen is placed in a magnetic field, the screening current which circulates to cancel the externally applied magnetic flux inside (Meissner effect) must flow within this surface layer. Consequently, the magnetic flux density does not fall abruptly to zero at the boundary of the superconducting specimen but decreases exponentially within the surface region within which screening the current is circulating.

In 1935, F. London and H. London applied the conditions $E = 0$ (zero resistivity) and $B = 0$ (Meissner effect) to Maxwell's electromagnetic equations. They showed that within a superconductor the applied field does not drop suddenly to zero at the surface of the superconductor, but decays exponentially according to the relation

$$B = B_o \exp\left(-\frac{x}{\lambda}\right) \quad \dots(8.6)$$

where B_o is the flux density at the surface and λ is the characteristics length called *penetration depth*. Using the above expression, we can define the penetration depth as the distance within the superconducting material from the surface where the flux density B_o falls to $\frac{B_o}{e}$. The variation of flux density inside a superconductor is shown in Fig. 8.10.

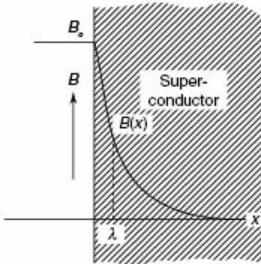


Fig. 8.10 Variation of magnetic flux density inside a superconductor.

For all practical purposes, the penetration depth (≈ 10 to 100 nm) is very small and for bulk superconductors the superconductor can be treated as perfectly diamagnetic. However, for the superconductors in thin film form, the penetration is not negligible and it effects the superconducting behaviour.

Variation of Penetration Depth with Temperature: The value of penetration depth depends upon the temperature of superconductor, as shown in Fig. 8.11. At a temperature much lower than T_c , it is nearly independent of temperature and has a characteristic value λ_o for a particular material. However, as the temperature approaches T_c , the penetration depth increases rapidly and approaches infinity. The variation of penetration depth with temperature is found to follow the relation

$$\lambda = \frac{\lambda_o}{(1-t^4)} \quad \dots(8.7)$$

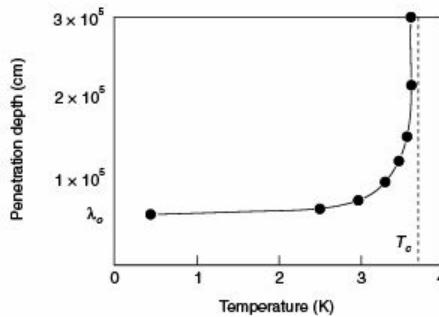


Fig. 8.11 Variation of penetration depth with temperature (in cm).

where, $t = \frac{T}{T_c}$. The penetration depth λ_o at 0 K for Hg, Pb, Al and In is 70 nm, 39 nm, 50 nm and 64 nm respectively.

8.8 THERMAL PROPERTIES

It is observed that the thermal properties like entropy, specific heat, etc. show sharp change on the material undergoing normal superconducting phase transition. As the Meissner effect shows that the transition in the presence of magnetic field is reversible, this implies that it is possible to apply the laws of thermodynamics to the normal superconducting phase transitions.

8.8.1 Entropy

The entropy of the superconductors decreases sharply as it is cooled below its transition temperature from its normal phase. As the entrophy is a measure of the disorderliness in the system, this implies that the *superconducting state is more ordered than the normal state*.

Figure 8.12 shows the variation of entropy of aluminium. The change in entropy is small and for Al the change is of the order of $10^{-4} k_B$ per atom. The small entropy difference suggests that the

rearrangement of the system on becoming superconducting is small. Measurements of gyromagnetic ratio suggests that in this rearrangement of conduction electrons are involved. These facts suggest that some of the conduction electrons, thermally excited in the normal state, get ordered in the superconducting state.

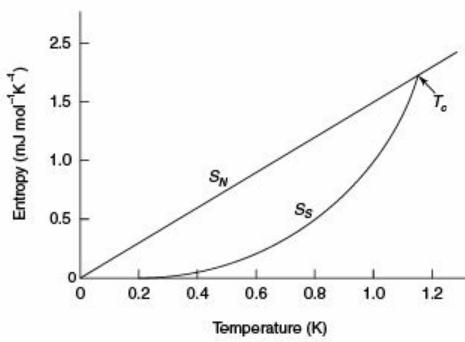


Fig. 8.12 Variation of entropy with temp. in normal (S_N) and superconducting (S_S) phases (for aluminium).

8.8.2 Specific Heat

The specific heat in the normal state of a superconducting material in the low temperature region can be expressed as

$$C_n = r T + \alpha T^3 \quad \dots(8.8)$$

where the first term represents the electronic contribution and the second term due to lattice vibrations (also known as Debye T^3 term). As shown in Fig. 8.13, the specific heat of a superconductor shows a jump at T_c . Because the onset of the superconductivity primarily involves electron, it is believed that the lattice vibration component to specific heat remains unaffected and only the electronic component changes. The experimental variation suggests that the electronic contribution to the heat capacity in the superconducting state is of exponential form

$$C_{se} = A \exp\left(-\frac{\Delta}{K_B T}\right) \quad \dots(8.9)$$

where A is a constant and 2Δ is an energy gap. The above expression is suggestive of excitation of electrons across a energy gap. The existence of an energy gap is a characteristic feature of superconducting state. The Bardeen-Cooper-

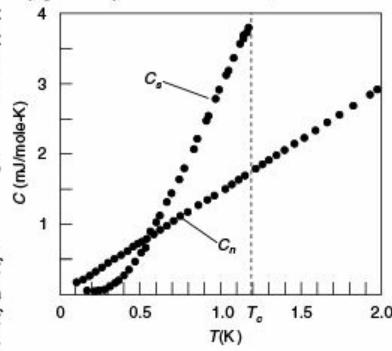


Fig. 8.13 Variation of specific heat in its normal (C_n) and superconducting (C_s) phase (for aluminium).

Schrieffer (BCS) theory of superconductivity (discussed later) explained the existence of the energy gap. The cooper pairs, the pairs of electrons responsible for superconductivity, have zero energy in the ground state and do not contribute to the heat capacity.

8.8.3 Thermal Conductivity

It is observed that thermal conductivity of an ideal superconductor drops markedly when superconductivity sets in. This suggests that the electronic contribution to the thermal conductivity drops in the superconducting state. Figure 8.14 shows the thermal conductivity results obtained for tin revealing the drop in thermal conductivity when superconductivity sets in. For tin ($T_c = 3.73$ K), the thermal conductivity reduces from $55 \text{ W cm}^{-1} \text{ K}^{-1}$ at 4 K to $16 \text{ W cm}^{-1} \text{ K}^{-1}$ at 2 K in the superconducting phase while it is $34 \text{ W cm}^{-1} \text{ K}^{-1}$ in the non-superconducting phase at the same temperature.

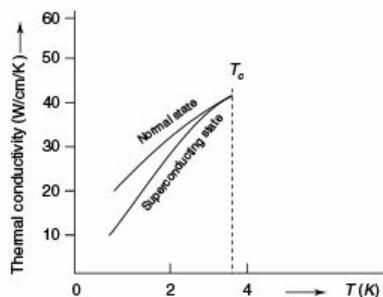


Fig. 8.14 Variation of thermal conductivity (of tin) with temp in normal and superconducting states.

The observed change is understood by the fact that superconducting electrons (or cooper pairs) play no part in heat transfer. In fact, this observation was the indication that one-electron (or independent electron) explanation mostly employed for the normal conductors is not applicable to superconductors. Usually, good conductors of electricity are good conductors of heat. In contrast, superconductors show low values of thermal conductivity in spite of showing extremely high electrical conductivity. This prompted scientists to believe that electrical current in a superconductor is not due to the flow of a single electron but due to the flow of electron pairs. The electron pairs have zero energy and cannot transport thermal energy, although, due to their net charge their flow constitutes an electric current. The difference between normal state thermal conductivity and that in superconducting state increases as the temperature reduces below the superconducting transition temperature as very few normal electrons are left in the material with the reducing temperature. Unlike in normal conductors, an electric current at a uniform temperature in a superconductor is not accompanied by a thermal current. This also explains the absence of Peltier effect in superconductors, implying that the superconducting electrons do not carry heat.

However, on the destruction of superconductivity by the application of a magnetic field ($H > H_c$), the thermal conductivity of the material increases. Due to this effect, superconductors have been utilized as *thermal switches* at low temperatures by making or breaking heat contact between specimens connected by a link of superconducting material.

8.8.4 Energy Gap

We have seen that the heat capacity in the superconducting state varies with the temperature in an exponential manner, i.e., $\exp\left(-\frac{\Delta}{k_B T}\right)$. Optical and tunnelling experiments have revealed that the

argument of the exponential factor is $-\frac{E_g}{2k_B T}$. An exponential dependence implies that it requires a finite energy Δ to excite an individual electron in a superconductor, indicating the existence of an energy gap in the superconductors. The energy gap has been found to be of the order of 10^{-4} eV. If this energy gap is expressed in the form of $k_B T$, T comes out to be around 1°C , of the order of T_c .

The energy band diagram of a superconductor is shown in Fig. 8.15. The energy gap in superconductors differ from that in semiconductors and insulators. In superconductors, the energy gap separates superconducting electronic states from the normal states lying above it. In semiconductor the energy gap is a short of barrier for the conduction electrons and this energy must be supplied to the electrons in the valence band to make them free so as to conduct the electricity. While in a superconductor the energy gap do not obstruct the current flow and do not affect the flow of electrons (or cooper pairs) carrying current in a superconductor. *In a superconductor, the gap shows the energy separation between the normal electrons and the paired electrons called cooper pairs.* It has been found that

$$E_g = 2 \Delta = 2 b (k_B T)$$

or

$$\frac{E_g}{k_B T} = 2b \quad \dots(8.10)$$

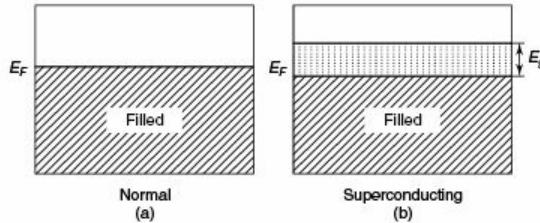
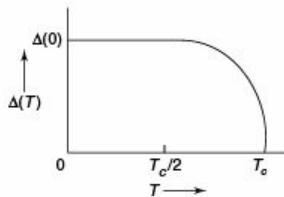


Fig. 8.15 Energy band diagram for (a) normal metal and (b) superconductor ($E_g \sim 10^{-4} E_F$)

The quantity Δ is often called the *energy gap parameter*. The energy gap decreases continuously to zero as the temperature is increased (Fig. 8.16) to the transition temperature T_c . The value of $2 b$ is about 3.5, i.e., the gap decreases from a value of about $3.5 k_B T$ at 0 K to zero at T_c . The value of energy gap of some of the elemental superconductors is given in Table 8.3.

Table 8.3 Energy gap 0 K

Element	$E_g(0)$ (in 10^{-4} eV) 2 $\Delta(0)$	$T_c(K)$	$\frac{E_g(0)}{k_B T_c}$
Ta	14.0	4.48	3.6
Al	3.4	1.2	3.3
In	10.5	3.4	3.6
Nb	30.5	9.5	3.8
Sn	11.5	3.72	3.6
Hg	16.5	4.18	4.6
Pb	27.3	7.18	4.4

**Fig. 8.16** Variation of energy gap with temperature.

Phase transitions are classified as the first order or second order depending upon the presence or absence of a latent heat and discontinuity in the energy gap. In a superconductor, the transition in zero magnetic field, from superconducting to normal state, is a second order phase transition as there is no latent heat but there is a discontinuity in the heat capacity.

Although, energy gap is an important parameter for superconductors, this is not a universal feature of superconductivity. There are superconducting materials without any evidence of energy gap. These gapless superconductors are, however, not typical and most superconducting materials exhibit a well-defined energy gap.

8.9 ABSORPTION OF ELECTROMAGNETIC RADIATION

The existence of an energy gap (2Δ) determines the frequency (ω) or energy ($\hbar \omega$) of the photon which is absorbed in a superconductor when an electromagnetic radiation is incident on it. The photons with energy less than 2Δ ($\hbar \omega < 2\Delta$) are not absorbed and nearly all the incident photons are reflected, i.e., a superconductor has zero resistance as determined by measuring reflection coefficient. For photons with energy less than the energy gap, the resistivity of a superconductor vanishes at absolute zero. At $T \ll T_c$, the resistance in the superconducting state has a threshold at the gap energy. Photons of lower energy see a resistanceless surface. However, the photons with energy higher than 2Δ cause excitation to the unoccupied normal energy states above the gap and the resistance approaches its normal value due to the absorption of photons. As the temperature is increased not only does the gap energy decreases but the resistivity for photons, with energy less than the gap energy, no longer vanishes, except at zero frequency. At zero frequency, the superconducting electrons short-circuit normal electrons which have been thermally excited above the gap. At finite frequencies, the inertia of superconducting electrons prevents them from completely screening the electric field. This allows the electric field to interact with thermally excited normal

electrons, leading to the absorption of incident electromagnetic radiation by the excited normal electrons.

As shown in Fig. 8.17, as the temperature is increased the resistance increases initially gradually and then very sharply as T_c is approached. For the photons with energy greater than the gap energy, the resistance is never zero and it approaches the resistance of the normal state at infrared frequencies. The transition from low frequency behaviour to high frequency behaviour occurs gradually. As the temperature is increased, the resistance increases sharply towards normal value.

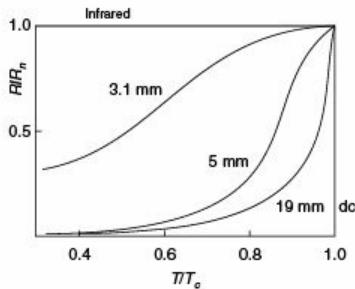


Fig. 8.17 $\frac{R}{R_n}$ (Ratio of resistance in the superconducting to normal state) as a function of $\frac{T}{T_c}$ in microwave region (for Al). In the infrared region superconducting and normal states are not distinguishable.

8.10 ISOTOPE EFFECT

In 1950, it was discovered by Maxwell and independently by Reynolds that *superconducting transition temperature of a material depends upon its isotopic mass*. They found that T_c for mercury varies from 4.186 K to 4.146 K as isotopic mass M of mercury change from 199.5 amu to 203.4 amu. Later this effect was also found in other elemental superconductors. It was found from the experimental results that

$$\text{or } M^\alpha T_c = \text{const} \quad \dots(8.11)$$

For most of the specimen α is observed to be around 0.5. The above relation is also expressed as

$$\frac{T_c}{\theta_D} = \text{const.} \quad \dots(8.12)$$

where θ_D is the Debye temperature. Equation (8.12) suggests that lattice vibration must have something to do with superconductivity. This result was very decisive for the development of the theory of superconductivity.

Although the change in T_c may not be drastic, the discovery of this effect was highly significant in the development of theoretical understanding of superconductivity. The dependence of T_c on the isotopic mass indicated that lattice vibrations and hence electron-lattice interactions are involved in superconductivity.

Example 8.4

The transition temperature of a superconducting specimen of isotopic mass 204.87 is 19.2 K. What will be the T_c with an isotopic mass of 218.87?

Solution.

$$\text{We know } T_c m^{1/2} = \text{constant}$$

$$\text{Here, } m = 204.87, T_c = 19.2 \text{ K}$$

$$m' = 218.87, T'_c = ?$$

$$\therefore (204.87)^{1/2} \times (19.2) = (218.87)^{1/2} T'_c$$

$$T'_c = \frac{\sqrt{204.87} \times 19.2}{\sqrt{218.87}} = 18.58 \text{ K}$$

8.11 LONDON EQUATION

The equation of motion of the superconducting electron is

$$m \frac{d\mathbf{v}_s}{dt} = e \mathbf{E}$$

$$\text{or } m \dot{\mathbf{v}}_s = e \mathbf{E} \quad \dots(8.13)$$

where \mathbf{v}_s is the velocity of the superelectrons (electrons responsible for superconductivity) and m and e are their mass and charge respectively. If n_s is the number of superelectrons per unit volume, then super current density is

$$\mathbf{J}_s = e n_s \mathbf{v}_s$$

$$\text{and } \frac{d\mathbf{J}_s}{dt} = e n_s \left(\frac{d\mathbf{v}_s}{dt} \right)$$

$$\text{or } \dot{\mathbf{J}}_s = e n_s \dot{\mathbf{v}}_s = \frac{n_s e^2 \mathbf{E}}{m} \quad \dots(8.14)$$

The above equation shows that an electric field produces a continuously increasing current at a rate determined by the above relation. This is the **first London equation**. Since no electric field is necessary unless the current changes, it is possible in a superconductor to have a steady current even in the absence of electric field.

The corresponding equation for the normal current is

$$\mathbf{J}_n = \sigma_n \mathbf{E} = \frac{n_n e^2 \tau \mathbf{E}}{m} \quad \dots(8.15)$$

Obviously, if $\mathbf{E} = 0$, $\mathbf{J}_n = 0$, i.e., there is no normal current density in the absence of an electric field. An electric field is necessary for the existence of a steady current.

According to Maxwell's equations, the magnetic field and the electric field are related by

$$\dot{\mathbf{B}} = -\operatorname{curl} \mathbf{E} \quad \dots(8.16)$$

$$\text{and } \operatorname{curl} \mathbf{H} = \mathbf{J} + \dot{\mathbf{D}} \quad \dots(8.17)$$

Here, we wish to obtain relations which relate the fields in superconductors to the currents flowing through them. We can express Eq. (8.17) as

$$\text{curl } \mathbf{B} = \mu_0 (\mathbf{J}_s + \dot{\mathbf{D}}) \quad \dots(8.18)$$

The displacement current $\dot{\mathbf{D}}$ is negligible in comparison to \mathbf{J}_s , unless the fields are varying very rapidly in time. Therefore, the above equation(s) can be simplified to

$$\dot{\mathbf{B}} = -\text{curl } \mathbf{E} \quad \dots(8.19)$$

and $\text{curl } \dot{\mathbf{B}} = \mu_0 \mathbf{J}_s \quad \dots(8.20)$

Substituting Eq. (8.14) in Eq. (8.19), we have

$$\dot{\mathbf{B}} = -\left(\frac{m}{n_s e^2}\right) \text{curl } \mathbf{j}$$

Eliminating \mathbf{J}_s using Eq. (8.18), we have

$$\dot{\mathbf{B}} = -\left(\frac{m}{\mu_0 n_s e^2}\right) \text{curl } \mathbf{j} \quad \dots(8.21)$$

For simplicity, using a single symbol α for the constant $\frac{m}{\mu_0 n_s e^2}$, the above equation can be

expressed as $\dot{\mathbf{B}} = \alpha \text{curl curl } \dot{\mathbf{B}}$

Using the identity, $\text{curl } \dot{\mathbf{B}} = \text{grad div } \dot{\mathbf{B}} - \nabla^2 \dot{\mathbf{B}}$ but for Maxwell's equations $\text{div } \dot{\mathbf{B}} = 0$, therefore,

$$\dot{\mathbf{B}} = \alpha \nabla^2 \dot{\mathbf{B}}$$

or $\nabla^2 \dot{\mathbf{B}} = \frac{1}{\alpha} \dot{\mathbf{B}} \quad \dots(8.22)$

This is a differential equation which \mathbf{B} must satisfy. We have deduced the above expression by applying the usual law of electrodynamics to a conductor with zero resistance. Therefore, the treatment is completely applicable to a perfect conductor but it is not so for a superconductor. We have seen from the Meissner effect that inside a superconductor the flux density is not only constant but the value of this constant is always zero, i.e., not only $\dot{\mathbf{B}}$ but $\dot{\mathbf{B}}$ itself must die away rapidly below the surface. London brothers suggested that the magnetic behaviour of a superconducting specimen might be correctly described if Eq. (8.22) is applied not only to $\dot{\mathbf{B}}$ but to \mathbf{B} itself. That is,

$$\nabla^2 \mathbf{B} = \frac{\mathbf{B}}{\alpha} \quad \dots(8.23)$$

The solution of this equation is of the form

$$B(x) = B_o \exp\left(-\frac{x}{\sqrt{\alpha}}\right)$$

If we trace the argument in arriving at Eq. (8.22), then we find that we could have derived Eq. (8.23) if everywhere we had replaced $\dot{\mathbf{B}}$ by \mathbf{B} and obtained the relation

$$\mathbf{B} = -\frac{m}{n_s e^2} \operatorname{curl} \mathbf{J}_s \quad \dots(8.24)$$

This equations along with the relation

$$\mathbf{j} = \frac{n_s e^2}{m} \mathbf{E} \quad \dots(8.25)$$

together describe the electrodynamics of the supercurrent and are known as *London equations*. Equation (8.24) describes the diamagnetism shown by the superconductors and Eq. (8.25) describes the resistanceless property of a superconductor; there being no electric field in the specimen unless the current is changing. London equations are obtained from general equations of electromagnetism so as to explain the experimentally observed behaviour of the superconductors. They do not explain the cause of the occurrence of superconductivity.

For a one-dimensional case, Eq. (8.23) takes the form

$$\frac{\partial^2 B(x)}{\partial x^2} = \frac{1}{\alpha} B(x) \quad \dots(8.26)$$

where B_a is the flux density of the applied magnetic at the surface. Equation (8.26) shows that the flux density dies away exponentially inside a superconductor, falling to $\left(\frac{1}{e}\right)$ of its value at the surface at a distance $x = \sqrt{\alpha}$ (Fig. 8.10). This distance is called the *London penetration depth*, λ_L .

As,

$$\alpha = \frac{m}{\mu_0 n_s e^2},$$

$$\therefore \text{London penetration depth, } \lambda_L = \sqrt{\frac{m}{\mu_0 n_s e^2}} \quad \dots(8.27)$$

The substitution of the values of m and e (electronic mass and charge) and $n_s = 4 \times 10^{28} \text{ m}^{-3}$ (the normal electronic concentration in metals, about one conduction electron per atom), λ_L is found to be of the order of 10^{-6} cm .

Equation (8.27) shows that the penetration depth λ is inversely proportional to $n_s^{1/2}$. As the concentration of superelectrons decreases with increasing temperature, the penetration depth increases with increasing temperature, rising to infinity as the temperature approaches the transition temperature (Fig. 8.11). The observation can well be approximated by the expression

$$\frac{\lambda(t)}{\lambda(0)} = \left[1 - \left(\frac{T}{T_c} \right)^4 \right]^{\frac{1}{2}} \quad \dots(8.28)$$

This implies that the density of superconducting electrons increases from zero at T_c to n_o at absolute zero and can be expressed as

$$n_s = n_o \left[1 - \left(\frac{T}{T_c} \right)^4 \right] \quad \dots(8.29)$$

The London equations can be expressed as

$$\operatorname{curl} \mathbf{J}_s = -\frac{1}{\mu_o \lambda_L^2} \mathbf{B} \quad \dots(8.30)$$

and $\mathbf{j}_s = \frac{1}{\mu_o \lambda_L^2} \mathbf{E} \quad \dots(8.31)$

The London equations are additional conditions obeyed by supercurrents. In the most general case, the total current density \mathbf{J} is the sum of the normal current \mathbf{J}_n and supercurrent density \mathbf{J}_s ,

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_s$$

The normal current obeys Maxwell's equations and Ohm's law

$$\mathbf{J}_n = \sigma \mathbf{E}$$

where σ_n is the conductivity associated with the normal electrons. For a superconducting specimen

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_s$$

$$\mathbf{J}_n = \sigma_n \mathbf{E}$$

$$\operatorname{curl} \mathbf{J}_s = -\frac{1}{\mu_o \lambda_L^2} \mathbf{B}$$

$$\mathbf{j}_s = \frac{1}{\mu_o \lambda_L^2} \mathbf{E} \quad \dots(8.32)$$

In the steady state when fields and currents are not changing with time, the only current is the supercurrent, i.e., $\mathbf{J}_n = 0$ and only London equations are applicable to a superconducting specimen, leading to the condition.

$$\nabla^2 \mathbf{B} = \frac{1}{\lambda_L^2} \mathbf{B} \quad \dots(8.33)$$

The above condition was a guess, which was fairly successful in explaining the experimentally observed behaviour of superconductors.

Example 8.5

Estimate the value of London penetration depth at 0 K for Pb if its density is $11.3 \times 10^3 \text{ kg/m}^3$ and its atomic weight is 207.19. Calculate the increase in penetration at 5 K from its value at 0 K if its T_c is 7.22 K.

Solution.

$$\text{London penetration depth } \lambda_o \text{ (at 0 K)} = \sqrt{\frac{m}{\mu_o n_s e^2}}$$

$$n = n_s = \frac{N_A \rho}{M_A} \quad (N_A = \text{Avogadro's number})$$

density $\rho = 11.3 \times 10^3 \text{ kg/m}^3, M_A = 207.19$
 $N_A = 6.023 \times 10^{26}/\text{Kmol}$

$$\therefore n_s = \frac{6.023 \times 10^{26} \times 11.3 \times 10^3}{207.19} = 3.28 \times 10^{28}/\text{m}^3$$

Also,

$$m = 9.1 \times 10^{-31} \text{ kg}, \quad e = 1.6 \times 10^{-19} \text{ C}$$

$$\mu_0 = 4\pi \times 10^{-7} \text{ henry/m}$$

$$\therefore \lambda_o = \sqrt{\frac{9.1 \times 10^{-31}}{4\pi \times 10^{-7} \times 3.28 \times 10^{28} \times (1.6 \times 10^{-19})^2}} \\ = 29.3 \text{ nm}$$

We know,

$$\lambda_T = \frac{\lambda_o}{\sqrt{1 - \left(\frac{T}{T_c}\right)^4}}$$

$$T = 5 \text{ K} \quad T_c = 7.22 \text{ K}$$

$$\lambda_T = \frac{29.38 \text{ mm}}{\sqrt{1 - \left(\frac{5}{7.22}\right)^4}}$$

$$= 33.4 \text{ nm}$$

Hence, increase in penetration depth = $33.4 - 29.3 = 4.1 \text{ nm}$

$$\text{Percentage increase } \frac{4.1}{29.3} \times 100 = 13.99\%$$

8.12 BCS THEORY

Since its discovery in 1911, despite intensive efforts, the origin of the superconductivity could not be understood for a long period of time. Finally in 1957, Bardeen, Cooper and Schrieffer developed a theory, known as BCS theory, which could explain superconductivity in the materials now known as low- T_c superconductors. The isotope effect, variation of specific heat, existence of energy gap and long range electronic order gave the indication that electrons in the superconductors are bound and the involvement of the lattice in the superconductivity phenomenon. In 1950, Herbet Frohlich developed the idea but he could not explain superconductivity. He had suggested that superconductivity arises due to phonon (quantum of lattice vibration) mediated electron-electron interaction. Based on this, J. Bardeen, L.N. Cooper and J. R. Schrieffer developed a theoretical picture of superconductivity.

The existence of the long-range order indicated that the electrons must interact with each other. The experimental evidence led to believe that the superconductivity is due to a cooperative behaviour of conduction electrons. However, any kind of cooperation or coherence between the electrons appeared unlikely for the two reasons; statistical and dynamical. First, electrons are fermions, obeying Fermi-Dirac statistics, so there cannot be more than one electron per state. On the other hand, coherence presumably would result if there is large number of particles in a single quantum state. Secondly, electrons repel each other due to the Coulombic force, which should make the electrons move away from each other. It was shown that both the hurdles are overcome by the electrons and

they effectively attract one another in certain energy range. Due to the attraction, they pair up and a pair of electron is referred to as **cooper pair**. An electron pair acts like a boson and so the pairs can all condense into a single quantum state.

The motion of the electrons may be represented by waves and in a perfect crystal at absolute zero, free from defects and thermal vibrations, these waves can propagate freely through the lattice without any attenuation. However, at any finite temperature, there are thermal vibrations and there is finite possibility of the wave being reflected or scattered. In other words, it can be said that electrons interact with the lattice or we can say that there is electron-lattice interactions. It is this electron-lattice interaction which determines the resistivity of metals and semiconductors at room temperature. Since both momentum and energy must be conserved when an electron is scattered, one of the vibrational modes of the lattice must be excited in the scattering process. This vibrational motion is quantized, and is expressed in terms of absorption (or emission) of a phonon. Thus, *a phonon is the same to lattice vibration as a photon is to light wave*.

Frohlich, in 1952, tried to explain the attraction between the electrons via lattice distortion. He pointed that the electron-phonon interaction was able to couple two electrons in such a way that they behaved as if there was a direct interaction between them. In this postulated interaction, one electron emits a phonon which is immediately absorbed by another and he was able to show that in certain circumstances, this emission and subsequent absorption of a phonon gives rise to a weak interaction between the electrons, which is capable of producing an energy gap of the right order of magnitude.

8.12.1 Cooper-Pair formation

Cooper extended Frohlich's idea on electron-phonon interaction into electron-phonon-electron interaction. Cooper demonstrated that with the creation of condition favourable for a net attractive interaction between two electrons in a conductor, the conductor is transferred from its normal conducting state to its superconducting state. The current in a superconductor is attributed to the flow of these cooper pairs. At absolute zero, all the electrons in states near the Fermi surface are bound into cooper pairs and are in the ground state. The existence of cooper pairs is manifested by the experimental evidence of flux quantization, discussed later in this chapter.

The interaction can be understood in the following way. Suppose an electron approaches a positive ion core. It suffers coulombic attractive interaction. Due to this attractive interaction, the ion core is set into motion and consequently distorts the lattice. Smaller the mass of the ion core, the greater is this distortion. Suppose another electron approaches this distorted lattice and interacts with it. The effect of this interaction is to lower the energy of the second electron. In other words, the two electrons interact via lattice distortion resulting in the lowering of energy for the electrons. The lowering of the energy implies that the force between the two electrons is attractive. This electron-lattice-electron interaction is strongest when the two electrons have equal and opposite momenta and spin.

The formation of cooper pair or the phonon mediated electron-electron interaction is shown in Fig. 8.18. Let the two electrons having initial wave vectors \mathbf{k}_1 \mathbf{k}_2 interact through phonon exchange. If \mathbf{k}'_1 and \mathbf{k}'_2 are their respective final wave vectors after the interaction, then the momentum conservation requires.

$$\mathbf{k}_1 - \mathbf{k}'_1 = \mathbf{k}_2 - \mathbf{k}'_2 = \mathbf{k}_{ph} \quad \dots(8.34)$$

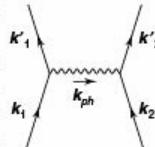


Fig. 8.18 Schematic representation of electron-electron interaction mediated by a phonon.

where k_{ph} is the wave vector of the phonon involved in the process. It is easily shown that the pairs with zero momentum (i.e., having equal and opposite momentum) are most likely to be formed. As shown in Fig. 8.19, consider two electrons with wave vectors \mathbf{k}_1 and \mathbf{k}_2 on the Fermi sphere such that $\mathbf{k}_1 = -\mathbf{k}_2 = \mathbf{k}_F$. The total momentum of the pair is $\mathbf{K} = \mathbf{k}_1 + \mathbf{k}_2$. When $\mathbf{K} \neq 0$, its magnitude is equal to $2k_F \cos \theta$ and only those electrons lying on the ring can form pairs. However, for $\mathbf{K} = 0$, i.e., $\mathbf{k}_1 = -\mathbf{k}_2$, all electrons on the Fermi surface can pair up because of the finite phonon energy $\sim \hbar$

$\omega_D = \frac{\hbar^2 k_F \Delta k}{M}$, $\hbar = \frac{h}{2\pi}$, ω_D – characteristic Debye frequency, cooper showed that lattice distortion can screen the coulomb repulsion between the two electrons resulting in net attraction if the energies of the two electrons is of the order of $\hbar \omega_D$. These electrons are confined in a thickness Δk of the shell on the fermi sphere.

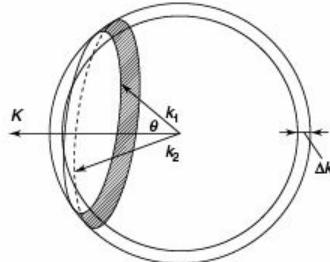


Fig. 8.19 A pair of electron states with wave vectors \mathbf{k}_1 and \mathbf{k}_2 on a shell of a spherical Fermi surface.
When $\mathbf{K} = \mathbf{k}_1 + \mathbf{k}_2 = 0$, all states can be paired.

Bardeen, Cooper and Schrieffer extended the above-mentioned Cooper's idea and developed a complete microscopic theory according to which favourable conditions for a condensation into bound pairs exist between electrons having equal and opposite wavevectors and opposing spins, i.e., $\mathbf{k} \uparrow$ and $-\mathbf{k} \downarrow$, keeping total momentum $\mathbf{K} = 0$.

Bardeen, Cooper and Schrieffer showed that if two electrons in a metal pair up to form a bound state, all the electrons in the metal tend to rearrange themselves into bound pairs. Since the characteristic dimension of each pair is larger ($\approx 4 \times 10^{-5}$ cm) than the average typical distance between two electrons ($\approx 10^{-8}$ cm), there is considerable overlap among the pairs. The coherence between them is thus of importance. They showed that the energy of the system is lowest when each pair has the same total momentum equal to zero. This is the single quantum mechanical state into which electron pairs condense.

8.12.2 Main Features of BCS Theory

Starting with the basic principles and using quantum mechanical treatment, the BCS theory was able to explain the effects shown by the superconductors, now referred to as low- T_c superconductors. The main features of the BCS theory can be summarized as follows.

- (i) The electron-lattice-electron interaction leads to attractive interaction and pair formation of electrons. When an electron interacts with the lattice, the deformation of lattice takes place. Another electron on finding the deformed lattice adjusts itself to lower its energy. Thus, the two electrons interact via lattice distortion.
- (ii) An attractive interaction between the electrons leads to a ground state separated from excited states by an energy gap. The critical magnetic field, thermal properties and most of the electromagnetic properties are related to the existence of the energy gap.
- (iii) In the presence of interactive interaction, the Fermi sea is unstable against the formation of bound pair states. In BCS Theory, simple form of potential is used, i.e., the potential is assumed to be constant over a narrow shell straddling the Fermi surface and is zero elsewhere. The width of the shell is $2\hbar\omega_D$.
- (iv) The penetration depth and coherence length emerge as natural consequences of the BCS theory. Also, the Meissner effect, the defining phenomenon in superconductivity follows naturally from this theory.
- (v) The paired states are situated within $k_B\theta_D$ of the Fermi energy, where θ_D is the Debye temperature.
- (vi) According to BCS theory, the transition temperature is estimated to be

$$T_c = 1.14 \theta_D \exp \left[-\frac{1}{UD(\epsilon_F)} \right] \quad \dots(8.35)$$

where, θ_D is the Debye temperature, U the attractive interaction and $D(\epsilon_F)$ the electron density of orbitals at the Fermi level. There is fairly good agreement with the experimental data.

- (vii) The magnetic flux through a superconducting ring is quantized. The effective unit of charge is found to be $2e$ and not e , confirming the pair formation.

8.12.3 BCS Ground State

For a non-superconducting material, the filled Fermi sea is the ground state of a Fermi gas of non-interacting electrons, i.e., it is filled Fermi sea bounded by the Fermi surface, (Fig. 8.20(a)). This state allows arbitrarily small excitations; by taking an electron from the Fermi surface and raising it just above the Fermi surface, an excited state can be obtained. While in a superconducting material, according to the BCS theory, there is an attractive interaction between the electrons. Therefore, in this case, we cannot obtain an excited state unless an energy exceeding the energy of attraction between the electrons is supplied. This implies that the BCS ground state is separated by a finite energy gap E_g

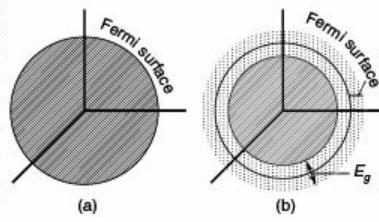


Fig. 8.20 (a) Ground state of the Fermi gas of non-interacting electrons. (b) The lowest excited state is separated from the ground state by an energy gap E_g .

from its lowest excited state. Further, since the interacting electrons are only those which are situated within $k_B \theta_D$ of the Fermi energy, this energy gap is situated about the Fermi surface of the Fermi state (Fig. 8.20(b)). In terms of the probability of occupation of the ground state of one particle state, the situation is depicted in Fig. 8.21(a) and (b) respectively for non-interacting and interacting electrons.

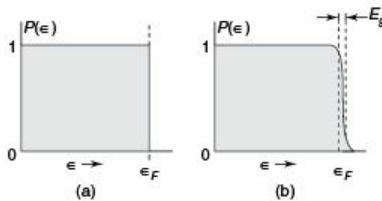


Fig. 8.21 (a) Probability P that an orbital of kinetic energy ϵ is occupied in the ground state of the non-interacting Fermi gas; (b) the BCS ground state differs from the Fermi state in a region of width of the order of the energy gap E_g (both curves are at 0 K).

Thus, the apparent features of the BCS ground state are as follows.

- The total energy of the BCS state is lower compared to Fermi state. It is because the total energy of the BCS state comprises the kinetic energy and the attractive potential energy, while the Fermi state comprises only the kinetic energy. This makes the BCS state more stable.
- The one particle states are occupied in pairs; if a state with wavevector $k\uparrow$ is occupied then the state with $-k\downarrow$ is also occupied. Similarly, if $k\uparrow$ is vacant, $-k\downarrow$ is also vacant.

For the quantum mechanical treatment, the knowledge of ground state wavefunction is very important. In a system of N electrons, the electrons are grouped into $N/2$ bound pairs with the state of each pair being described by a wavefunction $\psi(r_s)$, r being the position of electron and s the spin quantum number. The product of these $N/2$ identical two-electron wavefunctions is used to describe the N -electron wavefunction:

$$\phi(r_1 s_1, r_2 s_2, \dots, r_N s_N) = \psi(r_1 s_1, r_2 s_2) \dots, \psi(r_{N-1} s_{N-1}, r_N s_N) \quad \dots(8.36)$$

This represents a state in which all electrons are present as identical bound electron-pairs. Since the wavefunction ϕ represented by the above equation lacks the symmetry required by the Pauli principle, it must be antisymmetrized so that with the interchange of the space and spin coordinates of any two electrons the antisymmetrized wavefunction changes its sign. Then the wavefunction represents the BCS ground state:

$$\Phi_{BCS} = a \phi$$

where a is the antisymmetrizer the wavefunction Φ_{BCS} satisfies the Pauli principle, through it is constructed from identical pair function ψ . Unlike the case of single-electron states, the requirement of antisymmetry does not imply any restriction on the occupancy of two-electron levels in states which are antisymmetrized products of two-electron levels. It shows why it is possible for a pair of fermions to behave statistically like a boson, allowing all electron pairs to condense coherently into a single quantum state.

The complete quantum mechanical treatment of the BCS theory is beyond the scope of this book.

8.12.4 Coherence Length

According to the BCS theory, superconductivity arises due to the pairing of electrons or coherence of electrons. Coherence length is related to the distance an electron can feel the presence of another electron and can pair up. Coherence length, ξ is defined as the *mean separation at which electron-electron correlation becomes effective*. This is also regarded as the *average size of a cooper pair*.

Along with coherence length, London penetration depth λ_L is another important parameter often used for the classification of superconductors. In a pure superconductor, we define an intrinsic coherence length ξ_0 and according to BCS theory it is given by

$$\xi_0 = \frac{hv_F}{\pi E_g} \quad \dots(8.37)$$

where E_g is the energy gap and v_F is the electron velocity at the Fermi surface. In a pure type-I superconductor, the value of ξ_0 ranges from 1000 Å° to 10, 000 Å°. For example, ξ_0 (at 0 K) of Sn is 2300 Å°, for Al it is 16000 Å°, for Nb it is 380 Å°. For type II superconductors this parameter can be as low as 10 Å°. In impure superconducting specimen, the value of ξ is much shorter than ξ_0 .

8.13 MAGNETIC FLUX QUANTIZATION

Deaver and Fairbank in 1961 carried out experiments and precision measurements which showed that the magnetic flux through a superconducting ring (or a hollow superconducting cylinder) is

quantized in integral multiples of $\frac{h}{2e}$ ($= 2.07 \times 10^{-15}$ Weber). They electroplated a 5×10^{-4} cm

Thick tin cylinder on a copper wire of diameter 1.3×10^{-3} cm. The wire was cooled in the presence of a small magnetic field (~ 0.1 Gauss) to below the superconducting transition temperature of tin. The trapped flux was estimated by measuring the emf induced in an electrical coil when the cylinder was moved to and fro at a frequency of 100 Hz. The magnetic flux was found to be around the

values, $\frac{nh}{2e}$, ($n = 1, 2, 3, \dots$), within the experimental error.

Another team of scientists, led by Döll and Näbauer did similar experiments and measured the flux through the torque on the hollow superconducting cylinder in a known perpendicular magnetic field, also got the same result.

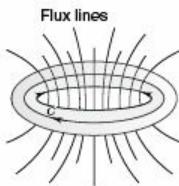


Fig. 8.22 A superconducting ring in a static magnetic field.

The measurements of magnetic flux have shown that $q = 2e$ confirming the formation of cooper pairs. Thus, any flux contained within a superconductor should exist only as multiples of a quantum, the **fluxon** (or **fluxoid**) ϕ_o given by

$$\phi_o = \frac{h}{2e} = 2.07 \times 10^{-15} \text{ Weber.}$$

8.14 TUNNELLING OF SINGLE PARTICLE

Quantum mechanics conceptualized the tunnelling of a particle across a barrier, long before its experimental observation. The tunnelling phenomenon was experimentally observed by Giaver in 1961.

As shown in Fig. 8.23 consider two metals A and B separated by a layer of insulator. The insulator acts as a barrier to the flow of conduction electrons from one metal to the other. If the insulating layer is thin ($\sim 10 \text{ \AA}$), there is a finite probability that an electron impinging on the barrier will pass through the barrier, this process is known as tunnelling.



Fig. 8.23 Metals A and B separated by insulating thin layer C.

When both the metals are normal conductors (even if one is superconductor but $T > T_c$), the current-voltage relation of the junction is ohmic, shown in Fig. 8.24(a). Giaver found that when one of the metal turns superconducting ($T < T_c$), the current voltage characteristic becomes curved as shown in Fig. 8.24(b) and not a straight line. The characteristic can be understood with the help of the energy diagram shown in Fig. 8.24(c). In the superconductor there is an energy gap at the Fermi level. If a small amount of voltage V is applied across the junction (electric field directed to left), the energy band of the left side is lowered by eV, (positive V applied to superconductor, lower its energy level, as the charge on the electron is negative) but the electrons are still unable to flow to the right due to unavailability of the unoccupied states on the right side due to the presence of the band gap. But if the voltage is raised further so that the energy band is down by $\Delta_o/2$, then the corresponding states on the right side are empty and the current starts to flow due to the tunnelling of electrons. The voltage at which current begins to flow is given by

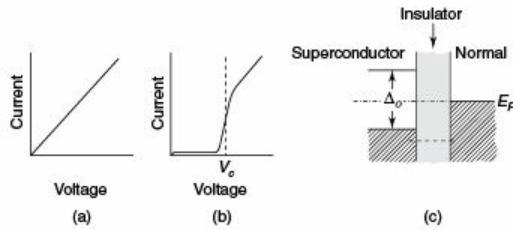


Fig. 8.24 (a) Metals A and B non-superconductors (b) One of the metal superconductors.

$$V = \frac{\Delta_a}{2e} \quad \dots(8.38)$$

The above expression is used to determine the superconducting gap this gap is essentially a measure of the energy needed to break an electron pair.

8.15 JOSEPHSON EFFECT-COOPER-PAIR TUNNELLING

In 1962, Brian D. Josephson theoretically predicted that the cooper pair can tunnel from one superconductor into another superconductor separated by a thin insulating layer also called a weak link (Fig. 8.25). This effect is known as *Josephson effect* or *Josephson tunnelling*. The insulating layer is very thin of the order of 10-20 Å. This effect is of immense significance and is due to the fact that the superconductor is characterized by a single wave function. Josephson effect is the manifestation of quantum interference phenomenon on a *macroscopic scale*. Josephson was awarded Nobel Prize in Physics in 1973.

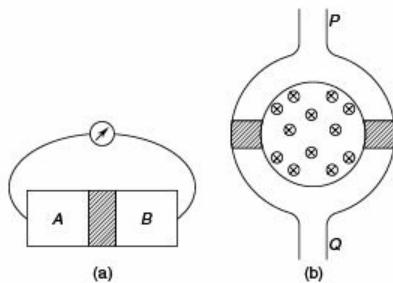


Fig. 8.25 (a) Two superconductors A and B with a thin insulating layer.
(b) Two superconductors P and Q with side arms and two weak links.

We discuss below the basic theoretical description of this effect. Depending upon whether a potential difference is applied across the superconductor-insulator-superconductor sandwich or not, there are two types of Josephson effect: d.c. Josephson effect (no potential or magnetic field applied) and a.c. Josephson effect (a d.c. voltage is applied).

When a d.c. magnetic field is applied through a superconducting circuit containing two junctions (Fig. 8.25 (b)), it causes maximum supercurrent to show *interference effects* as a function of magnetic field intensity (Fig. 8.2b). This effect is utilized for the design of SQUID (Superconducting Quantum Interferometer Device) which is used as ultrasensitive magnetometer, discussed later in this chapter.

8.15.1 D.C. Josephson Effect

In d.c. Josephson effect, a d.c. current flows across the junction in the absence of any applied electric or magnetic field. An ammeter connected across the junction shows a deflection even if no voltage is applied. We discuss here the theoretical description of this effect.

Let ψ_1 and ψ_2 be the wavefunctions of the cooper pairs on the two superconducting sides 1 and 2 respectively. For simplicity, let us assume that both the superconductors are identical and both are

at zero potential. The time dependent Schrödinger equation $i \hbar \frac{\partial \psi}{\partial t} = H\psi$ applied to the two sides gives

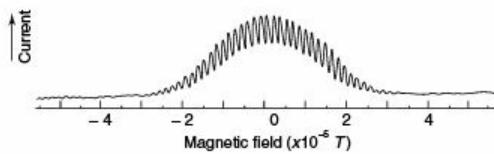


Fig. 8.26 Dependence of the supercurrent on the magnetic field in an SQUID.

$$i \hbar \frac{\partial \psi_1}{\partial t} = \hbar T \psi_2 \quad \dots(8.39)$$

and $i \hbar \frac{\partial \psi_2}{\partial t} = \hbar T \psi_1 \quad \dots(8.40)$

where $\hbar T$ represents the effect of the electron-pair coupling or transfer interaction across the insulator. T has the dimension of rate (or frequency) and is a measure of the leakage of ψ_1 into the region 2 and that of ψ_2 into the region 1. If the thickness of the insulator is too large there is no tunnelling of cooper pairs and T is zero.

Let $\psi_1 = n_1^{1/2} e^{i\theta_1}$ and $\psi_2 = n_2^{1/2} e^{i\theta_2}$ (n_1 and n_2 are the number density of cooper pairs on the two sides) and substituting these in Eq. (8.39), we get

$$\frac{\partial \psi_1}{\partial t} = \frac{1}{2} n_1^{-1/2} e^{i\theta_1} \frac{\partial n_1}{\partial t} + i n_1^{1/2} e^{i\theta_1} \frac{\partial \theta_1}{\partial t} = -iT \psi_2$$

or $\frac{\partial \psi_1}{\partial t} = \frac{1}{2} n_1^{-1/2} e^{i\theta_1} \frac{\partial n_1}{\partial t} + i n_1^{1/2} e^{i\theta_1} \frac{\partial \theta_1}{\partial t} = -iT n_2^{-1/2} e^{i\theta_2} \quad \dots(8.41)$

Similarly, $\frac{\partial \psi_2}{\partial t} = \frac{1}{2} n_2^{-1/2} e^{i\theta_2} \frac{\partial n_2}{\partial t} + i n_2^{1/2} e^{i\theta_2} \frac{\partial \theta_2}{\partial t} = -iT n_1^{1/2} e^{i\theta_1} \quad \dots(8.42)$

Multiplying Eq. (8.41) by $n_1^{1/2} e^{-i\theta_1}$ and using $\delta = \theta_2 - \theta_1$, we get

$$\frac{1}{2} \frac{\partial n_1}{\partial t} + i n_1 \frac{\partial \theta_1}{\partial t} = -iT (n_1 n_2)^{1/2} e^{i\delta} \quad \dots(8.43)$$

Similarly, multiplying Eq. (8.42) by $n_2^{1/2} e^{-i\theta_2}$, we get

$$\frac{1}{2} \frac{\partial n_2}{\partial t} + i n_2 \frac{\partial \theta_2}{\partial t} = -iT (n_1 n_2)^{1/2} e^{-i\delta} \quad \dots(8.44)$$

Equating real and imaginary parts of Eqs. (8.43) and (8.44), we have

$$\frac{\partial n_1}{\partial t} = 2T (n_1 n_2)^{1/2} \sin \delta; \quad \frac{\partial n_2}{\partial t} = -2T (n_1 n_2)^{1/2} \sin \delta \quad \dots(8.45)$$

and

$$\frac{\partial \theta_1}{\partial t} = -T \left(\frac{n_2}{n_1} \right)^{1/2} \cos \delta; \quad \frac{\partial \theta_2}{\partial t} = -T \left(\frac{n_1}{n_2} \right)^{1/2} \cos \delta \quad \dots(8.46)$$

As the two superconductors are assumed to be identical, taking $n_1 \equiv n_2$, from Eq. (8.46), we have

$$\frac{\partial \theta_1}{\partial t} = \frac{\partial \theta_2}{\partial t} \Rightarrow \frac{\partial(\theta_2 - \theta_1)}{\partial t} = 0 \quad \dots(8.47)$$

From Eq. (8.45), we find that

$$\frac{\partial n_2}{\partial t} = \frac{\partial n_1}{\partial t} \quad \dots(8.48)$$

The current flows across the junction, i.e., from superconductor 1 to superconductor 2 is proportional to $\frac{\partial n_2}{\partial t}$ or $-\frac{\partial n_1}{\partial t}$. We can, therefore, conclude from Eq. (8.48) that the supercurrent J across the junction due to the flow of the electron pairs depends on the phase difference and is expressed as

$$J = J_o \sin \delta = J_o \sin (\theta_2 - \theta_1) \quad \dots(8.49)$$

where J_o is proportional to the transfer interaction T . The current J_o is the maximum zero-voltage current that can pass through the junction. Without applying any voltage, a d.c. current will flow across the junction with a value between J_o and $-J_o$ depending upon the value of phase difference $\theta_2 - \theta_1$. This effect is known as *d.c. Josephson effect*.

8.15.2 A.C. Josephson Effect

Suppose a voltage V is applied across the junction. The junction being an insulator a potential can be applied across it. An electron pair experiences a potential energy difference qV ($q = -2e$) on passing across the junction. We can treat this as if a pair on one side is at a potential energy $-eV$ and a pair on other side is at eV . Then the equations of motion of the pair can be represented as:

$$i\hbar \frac{\partial \psi_1}{\partial t} = \hbar T \psi_2 - eV \psi_1 \quad \dots(8.50)$$

and

$$i\hbar \frac{\partial \psi_2}{\partial t} = \hbar T \psi_1 - eV \psi_2 \quad \dots(8.51)$$

Let

$$\psi_1 = n_1^{1/2} e^{i\theta_1} \text{ and } \psi_2 = n_2^{1/2} e^{i\theta_2}.$$

$$\therefore \frac{\partial \psi_1}{\partial t} = \frac{1}{2} n_1^{-1/2} e^{i\theta_1} \left(\frac{\partial n_1}{\partial t} \right) + n_1^{1/2} e^{i\theta_1} \frac{\partial \theta_1}{\partial t}$$

Substituting in Eq. (8.50), we get

$$i\hbar e^{i\theta_1} \left[\frac{1}{2} n_1^{-1/2} \frac{\partial n_1}{\partial t} + n_1^{1/2} i e^{i\theta_1} \frac{\partial \theta_1}{\partial t} \right] = \hbar T n_2^{1/2} e^{i\theta_2} - eV n_1^{1/2} e^{i\theta_1}$$

Dividing throughout by $i\hbar e^{i\theta_1} n_1^{-1/2}$, we get

$$= \frac{1}{2} \frac{\partial n_1}{\partial t} + i n_1 \frac{\partial \theta_1}{\partial t} = ie n_1 V \hbar^{-1} - i T(n_1 n_2)^{1/2} e^{i\delta} \quad \dots (8.52)$$

or $\frac{1}{2} \frac{\partial n_1}{\partial t} + i n_1 \frac{\partial \theta_1}{\partial t} = ie n_1 V \hbar^{-1} - i T(n_1 n_2)^{1/2} (\cos \delta + i \sin \delta)$

Equating real and imaginary parts, we have

$$\frac{1}{2} \frac{\partial n_1}{\partial t} = T(n_1 n_2)^{1/2} \sin \delta \quad \text{or} \quad \frac{\partial n_1}{\partial t} = 2T(n_1 n_2)^{1/2} \sin \delta \quad \dots (8.53)$$

and $\frac{\partial \theta_1}{\partial t} = \frac{eV}{\hbar} - T \left(\frac{n_2}{n_1} \right)^{1/2} \cos \delta \quad \dots (8.54)$

This differs from Eq. (8.46) by the presence of $\frac{eV}{\hbar}$.

Similarly, from Eq. (8.51), we can have

$$\frac{1}{2} \frac{\partial n_2}{\partial t} + i n_2 \frac{\partial \theta_2}{\partial t} = -ie V n_2 \hbar^{-1} - i T(n_1 n_2)^{1/2} e^{-i\delta}$$

which on separating real and imaginary parts yields

$$\frac{\partial n_2}{\partial t} = -2T(n_1 n_2)^{1/2} \sin \delta \quad \dots (8.55)$$

and $\frac{\partial \theta_2}{\partial t} = - \left(\frac{eV}{\hbar} \right) - T \left(\frac{n_1}{n_2} \right)^{1/2} \cos \delta \quad \dots (8.56)$

Taking $n_1 \approx n_2$, from Eqs. (8.55) and (8.56), we get

$$\frac{\partial(\theta_2 - \theta_1)}{\partial t} = \frac{\partial \delta}{\partial t} = -\frac{2eV}{\hbar} \quad \dots (8.57)$$

Integrating this expression, we find that with a d.c. voltage across the junction, the relative phase of the probability amplitudes vary as

$$\delta(t) = \delta(o) - \left(\frac{2eVt}{\hbar} \right) \quad \dots (8.58)$$

Therefore, with analogy to Eq. (8.49), the current is given by

$$J = J_o \sin \left[\delta(o) - \frac{2eVt}{\hbar} \right] \quad \dots (8.59)$$

This shows that the frequency of the resulting current is given by

$$\omega = \frac{2eV}{\hbar} \quad \dots (8.60)$$

This is the *a. c. Josephson effect*. If a voltage $V = 1 \mu V$ is applied across the junction, using Eq (8.60) it is seen that the frequency of current is 483.6 MHz. Equation (8.60) can also be interpreted as that a photon of energy $\hbar \omega = 2 eV$ is emitted or absorbed when an electron pair crosses the

barrier. By determining the voltage and frequency it is possible to obtain a precise value of $\frac{e}{\hbar}$, involving fundamental constants in physics.

8.16 MACROSCOPIC QUANTUM INTERFERENCE - SQUID

When a d. c. magnetic field is applied through a superconducting circuit (Fig. 8.22) containing two junctions, the maximum supercurrent shows interference effects as a function of magnetic field. Based on this effect, an extremely important device called *Superconducting Quantum Interference Device* (SQUID) has been designed. A SQUID is an *ultra sensitive magnetometer, capable of even detecting the change in magnetic field produced by the current pulse in human brain*. Therefore, it is used for the investigation of pathological changes in brain in addition to other applications based on detection of changes in even small magnetic field such as search of ore deposits, oil exploration, magnetic separation, earthquake research, submarine detection, etc.

For the observation of interference effect, two Josephson junctions are arranged in parallel, as shown in Fig. 8.27. A magnetic field is applied in the enclosed region normal to the plane of the combination. A supercurrent entering through the region 1 is divided into two parts and takes the two parallel paths, each of which contains a tunnel junctions. Let J_a and J_b be the supercurrent crossing the junctions a and b respectively. The two currents again combine in region 2. This combined current shows oscillations characteristic of an interference pattern produced by two coherent sources (Fig. 8.26). By analogy with interference of light, two currents J_a and J_b are treated as two coherent sources, whose superposition interference effect.

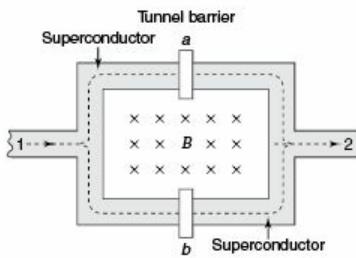


Fig. 8.27 Experimental arrangement for superconducting quantum interference.

A supercurrent flows between the two points if a phase difference (of the wavefunction) exists between the points (phase can be changed by application of magnetic field and voltage). If δ_a and δ_b are the phase shifts at the two barriers (in the absence of any applied magnetic field), the supercurrents through the two junctions can be expressed as

$$J_a = J_o \sin \delta_a \quad \dots(8.61)$$

$$J_b = J_o \sin \delta_b \quad \dots(8.62)$$

The phase shift in the presence of a magnetic field B and corresponding vector potential A can be expressed as

$$\nabla\theta = \frac{2e}{\hbar} A \quad \dots(8.63)$$

Taking a line integral of the above equation between two points P and Q , we have

$$|\Delta\theta|_P^Q = \int_P^Q \nabla\theta \cdot d\ell = \frac{2e}{\hbar} \int_P^Q A \cdot d\ell \quad \dots(8.64)$$

Therefore, the total phase shifts in the wavefunction along the two paths from region 1 to region 2 can be expressed as

$$|\Delta\theta|_1^2 = \delta_a + \frac{2e}{\hbar} \int_a A \cdot d\ell \quad \dots(8.65a)$$

and $|\Delta\theta|_1^2 = \delta_b - \frac{2e}{\hbar} \int_b A \cdot d\ell \quad \dots(8.65b)$

The two line integrals are in opposite direction, therefore, the negative sign in Eq. (8.65b)). Further, the two phase shifts must be identical because the wavefunction has a unique value at every point. Taking the phase shifts for the two parallel paths as identical and applying Stokes' theorem, we have

$$\delta_b - \delta_a = \frac{2e}{\hbar} \oint A \cdot d\ell = \frac{2e}{\hbar} \int B \cdot ds \quad \dots(8.66)$$

Taken together we get integral over a closed path. The above expression shows that the total phase difference is dependent on the value of the applied magnetic field. The general expression for δ_a and δ_b in Eq. (8.66), however, may be put as

$$\delta_a = \delta_o - \frac{e}{\hbar} \int B \cdot ds \quad \dots(8.67)$$

$$\delta_b = \delta_o + \frac{e}{\hbar} \int B \cdot ds \quad \dots(8.68)$$

where δ_o is an arbitrary phase shift whose value depends upon the nature of the barrier.

The total recombined current J can be expressed as

$$\begin{aligned} J &= J_a + J_b \\ &= J_o \left\{ \sin \left(\delta_o - \frac{e}{\hbar} \int B \cdot ds \right) + \sin \left(\delta_o + \frac{e}{\hbar} \int B \cdot ds \right) \right\} \\ &= 2 J_o \sin \delta_o \cos \left(\frac{e}{\hbar} \int B \cdot ds \right) \end{aligned} \quad \dots(8.69)$$

Thus, the cosine term gives the interference effect on the total current. The total current varies with $\int B \cdot ds$ and has maxima when

$$\frac{e}{\hbar} \int B \cdot ds = n\pi, n = 1, 2, 3, \dots$$

or

$$\phi = n \left(\frac{h}{2e} \right) \quad \dots(8.70)$$

$$= n \text{ (flux quantum)} \quad \dots(8.71)$$

The above expressions imply that a change of flux quantum in the enclosed flux change the current to a new maximum, as shown in Fig. 8.26 representing variation of current as a function of magnetic field.

8.17 HIGH-TEMPERATURE SUPERCONDUCTORS

Since the discovery of superconductivity in 1911 in Hg with $T_c = 4.2$ K, scientists have worked world over to find the materials with higher T_c . In 1953, superconductivity was found in Nb_3Ge with a $T_c = 23.2$ K. For over two decades it seemed that we would never cross this 23 K barrier. However, in a pleasant surprise, Bednorz and Muller in 1986 discovered superconductivity in $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ with a T_c close to 30 K. While the parent compound La_2CuO_4 is insulator, doping by alkaline earth materials, i.e., $\text{La}_{2-x}\text{M}_x\text{CuO}_{4-\delta}$ ($\text{M} = \text{Ba}, \text{Sr}, \text{Ca}$) induces superconductivity. For $\text{M} = \text{Sr}$ and $x = 0.15$, this compound gives the maximum T_c of 40 K. This discovery generated a renewed interest worldwide for finding new materials with higher T_c as discussed below. Bednorz and Muller were awarded Nobel Prize for their discovery.

Cuprates

$\text{La}_{2-x}\text{M}_x\text{CuO}_4$ has the quasi two-dimensional structure of K_2NiF_4 . Since the discovery of superconductivity in cuprates of K_2NiF_4 structure, many other cuprates with considerable higher T_c have been discovered and the common feature of them is the possession of Cu-O sheets. In Table 8.4 few families of cuprate superconductors are listed.

In 1987, superconductivity was discovered in $\text{YBa}_2\text{Cu}_3\text{O}_7$ (YBCO) with T_c upto 91 K making it possible to have a superconductor above the liquid nitrogen temperate (77 K). As liquid nitrogen is much cheaper than liquid helium, the coolant used earlier, it was a big breakthrough. The crystal structure of the YBCO system is shown in Fig. 8.28(a).

Table 8.4 Representative high- T_c superconductors and their T_c

Material	T_C (K)
1. $\text{La}_{2-x}\text{M}_x\text{CuO}_4$ $\text{M} = \text{Ca}, \text{Sr}, \text{Ba}$	20-40
2. $\text{LnBa}_2\text{Cu}_3\text{O}_7$ (123) $\text{Ln} = \text{Y}, \text{Eu}, \text{Pr Ce}, \dots$	91 K
3. $\text{Bi}_2\text{Sr}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+4}$ (BSCCO) $n = 1 (2201)$ $n = 2 (2212)$ $n = 3 (2223)$	20 85 110
4. $\text{Tl}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+4}$ (TBCCO) $n = 1 (2201)$ $n = 2 (2212)$ $n = 3 (2223)$	85 110 125

Contd.

Contd.

5. Hg Ba ₂ Ca _{n-1} Cu _n O _{2n+2}		
n = 1 (1212)		128
n = 2 (1223)		135
6. Sn Pb _{0.4} In _{0.6} Ba ₄ Tm ₅ Cu ₂ O ₂₀		181
7. Nd _{2-x} Ce _x Cu O ₄		25

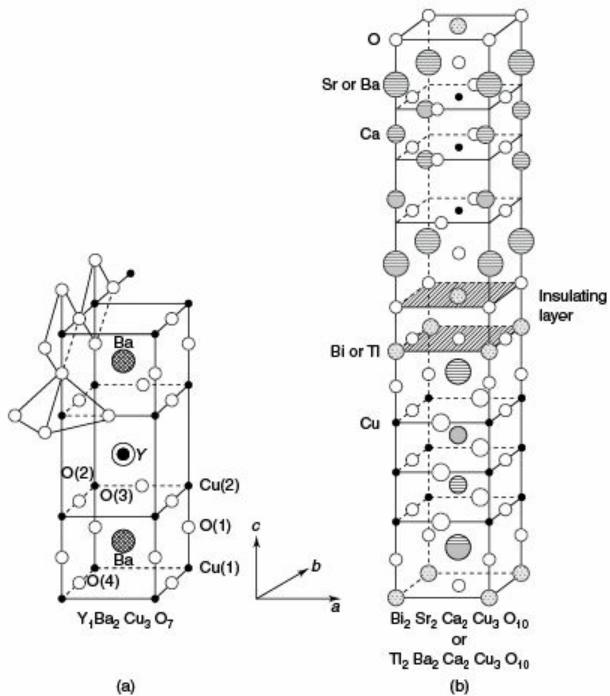


Fig. 8.28 (a) Unit cell structure of a $\text{Y}_1\text{Ba}_2\text{Cu}_3\text{O}_7$ crystal. The numbers in brackets represent the special sites of oxygen and copper atoms in CuO_2 layers. (b) Unit cell structure of $\text{Bi}_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$ or $\text{Tl}_2\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$ crystal.

It has an orthorhombic structure and has two Cu-O sheets in the ab plane and Cu-O chains along the b -axis. Superconductivity is highly sensitive to oxygen content in $\text{Y}_1\text{Ba}_2\text{Cu}_3\text{O}_{7-x}$ and gives best superconductivity for $x = 0$ to 0.1 with highest T_c of around 91 K. For $x \approx 0.6$. The superconducting disappears and the material behaves as an antiferromagnetic insulator.

The discovery of superconductivity in Bi-Sr-Cu-O system ($T_c \approx 20$ K) was the first step towards the development of the high- T_c superconducting materials without any rare earth element as one of the elements. Superconducting transition temperature was raised to over 100 K by adding Ca to the

system. There are two prominent superconducting phases in Bi-Sr-Ca-Cu-O system ($\text{Bi}_2\text{Sr}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+4}$, $n = 1, 2, 3$); The high- T_c 2223 ($\text{Bi}_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$) phase with a T_c around 110 K and the low- T_c 2212 ($\text{Bi}_2\text{Sr}_2\text{Ca}_1\text{Cu}_2\text{O}_8$) phase with T_c around 85 K. Thallium based superconducting system Tl-Ba-Ca-Cu-O ($\text{Tl}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+4}$, $n = 1$ to 4) has four distinct superconducting phases with T_c in the range of 85 K to 125 K. Mercury based cuprates $\text{Hg}\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+2}$ give maximum temperature of 135 K for $n = 3$ ($\text{Hg Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_8$). The T_c can be increased upto 165K under high pressure, which is close to the lowest temperature (183 K) recorded on earth.

In Table 8.4 are listed some of the representative high- T_c superconducting materials with their maximum T_c . All of them, except $\text{Nd}_{2-x}\text{Ce}_x\text{Cu O}_{4-y}$ (discovered in 1989) have holes as charge carriers. This was the first n -type high- T_c superconductor with electrons as charge carrier.

Numerous efforts have been made and are being made to develop new superconducting materials with still higher values of superconducting transition temperature. Reports by many group of researches of existence of superconductivity at much higher temperature but superconductivity in these materials lack reproducibility. These observations, known as USO (unconfirmed superconductivity observation) only indicates that the nature might have more secrets and some day, hopefully, we will have room temperature superconductors.

8.17.1 High- T_c Superconductors: Characteristics and Mechanism

Most of the known high-temperature superconductors (HTSC) possess the following features.

- (i) They are composites of more than three elements.
- (ii) They have Cu-O₂ two-dimensional planes in the crystal structure.
- (iii) They have low carrier density $\approx 10^{21} \text{ cm}^{-3}$.
- (iv) Holes are the charge carriers except in $\text{Nd}_{2-x}\text{Ce}_x\text{Cu O}_4$.
- (v) They show anisotropy due to the layered structure.

Superconductivity occurs by substitution of the constituent elements (e.g., $\text{La}_2\text{Cu O}_4 \rightarrow \text{La}_{2-x}\text{Sr}_x\text{Cu O}_4$) or by the change of oxygen content (e.g., $\text{YBa}_2\text{Cu}_3\text{O}_6 \rightarrow \text{YBa}_2\text{Cu}_3\text{O}_{6+x}$) These processes change the average valency of Cu ions and introduce carriers, holes in most of the cases, into Cu-O₂ planes. The valency of Cu is 2+ in the Cu-O₂ plane in the parent compounds such as $\text{La}_2\text{Cu O}_4$, $\text{Nd}_2\text{Cu O}_4$ or $\text{YBa}_2\text{Cu}_3\text{O}_6$. As a consequence, the parent compounds are insulators. A small amount of doping (i.e., $x = 0.01$ or less in $\text{La}_{2-x}\text{Sr}_x\text{Cu O}_4$) induces metallic conduction and superconductivity of changing the valency of Cu. The transition temperature T_c also depends on the number of Cu-O₂ layers in the unit cell. This is clearly demonstrated by $\text{Bi}_2\text{Sr}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+4}$ system; $T_c \approx 20$ K for $n = 1$ (single layer), $T_c \approx 85$ K for $n = 2$ (double layer) and $T_c \approx 110$ K for $n = 3$ (triple layer). Same is the case with Tl-cuprates. It is remarkable to observe that this increase of number of layers is effected only when Cu-O₂ layers are separated by Ca layer. Superconducting Cu-O₂ planes are separated by charge reservoir building blocks. In $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$, Cu-O chains act as charge reservoir. In Bi- and Tl-based cuprates Bi-O and Tl-O layers act as charge reservoirs.

The low carrier density and high- T_c result in a relatively larger value of magnetic penetration depth ($\approx 2000 \text{ \AA}$ in Cu-O₂ plane and 7000 \AA perpendicular to this plane), and extremely short coherence length ($\approx 10 \text{ \AA}$ to 20 \AA in Cu-O₂ plane and less than 10 \AA in perpendicular direction). The high- T_c superconductors are ceramic oxides and not metals. They are brittle and not ductile like metals.

Weak links at grain boundaries in a superconducting sample are responsible for relatively low critical current density J_c values, particularly in bulk samples.

High- T_c superconductors are characterized by two dimensionally confined carriers, extremely short coherence length, a reduced isotope effect and a transition temperature too high to be accounted for by phonon coupling of the BCS theory. On the other hand, quasi particles of charge $2e$ and temperature dependence of energy gap, pairing with a s -wave symmetry, reflect a BCS like character. It is believed that electron correlation plays an important role in the high- T_c systems, and the basic issue is of identifying the origin of the strong binding energy of the pairing interaction for high- T_c superconductivity. Most theorists believe that high- T_c phenomenon in cuprate oxides is a combination of phononic and electronic mechanism. The former is always there to a small measure and the latter leads to a considerable enhancement of T_c .

Conventional mechanism of pairing via electron-phonon cannot lead to a T_c greater than 30 K. This small value of T_c obtained from electron-phonon interaction is due to the fact that $T_c \leq 0.2 \theta_D$, where θ_D is the Debye temperature. Therefore, it is logical to look for the new mechanisms if higher transition temperatures are to be accounted for. According to BCS theory

$$k_B T_c \equiv 1.14 \frac{\hbar \omega_D}{N(E_F) V} \text{ ext} \left[-\frac{1}{N(E_F) V} \right] \quad \dots(8.72)$$

where ω_D is the characteristic Debye phonon frequency, $N(E_F)$ is the density of state per unit energy per unit volume at the Fermi surface and v ($|v| > 0$) is the electron-phonon mediated coupling constant. Hence, it is logical to blame low T_c on the low Debye energy $\hbar \omega_D$ of the phonons. This immediately suggests that a higher energy Boson, if present, might act in place of (or in addition to) phonon yielding a higher T_c .

Theoretical approach towards high- T_c superconductors can broadly be divided into two categories: (i) BCS like theories and (ii) non-BCS like theories. In the BCS like theories, instead of phonons, exchange of electronic excitations, e.g., excitons, plasmon, etc. are used as pairing mechanism. In this way, i.e., by replacing phonons by a exciton or plasmon, several attempts have been made to explain higher T_c of oxide superconductors. Spin-fluctuation mediated paring or magnons pairing mechanism has also been proposed.

Among non-BCS like theories, resonance valence band (RVB) model by Anderson and bi-polaronic model of Chakraverty are noteworthy. The RVB model describes the occurrence of superconductivity in these oxides to be due to mobility of pre-existing singlet pairs, which are frozen in Mott-insulating state but become mobile in the doped phase due to creation of vacancies. Bi-polaronic model proposed base condensation of certain real space pairs called bi-polarons or bisqueeps as the mechanism of high-temperature superconductivity.

Many other theories have been proposed. However, a comprehensive theory of high- T_c superconductivity is yet to be worked out, like BCS theory of conventional low- T_c superconductor. Thus, the mechanism of high- T_c superconductivity is yet to be understood completely.

8.18 SUPERCONDUCTING FULLERIDES

In 1991, a new class of superconductors were discovered. Superconductivity was found in fullerenes whose prominent members are C_{60} -based molecular crystals. As shown in Fig. 8.29, the structure of a signal C_{60} molecule consists of 60 carbon atoms. It is a cluster of carbon atoms arranged in the shape of a truncated icosahedron with 20 hexagonal and 12 pentagonal faces (similar to that in

benzene, graphite and other organic molecules). The pentagon occurs due to the topological requirement for producing a closed structure which resembles a football. The C_{60} clusters form the basis for a three-dimensional crystal structure of C_{60} which is characterized by a FCC unit cell. Other fullerenes have been obtained from C_{60} by crystallizing C_{60} with alkali metals whose atoms are located in the gap between C_{60} spheres. The T_c of these fullerides (Table 8.5) are high enough to put them among high- T_c superconductors.

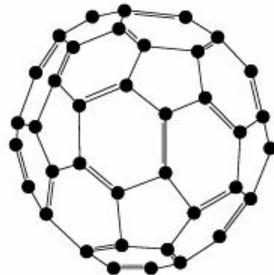


Fig. 8.29 The molecular of C_{60} .

Table 8.5 Some fullerides and their T_c

Fullerene	T_c (K)
$K_3 C_{60}$	19
$Rb_3 C_{60}$	28
$Rb Cs_2 C_{60}$	33
$Rb_{2.7} Tl_{0.2} C_{60}$	48

The superconducting fullerides, like ceramic high- T_c materials, show gradual onset of diamagnetism when cooled in zero magnetic field to temperatures below T_c . Due to the tendency to react with atmospheric oxygen, the fullerides have not been experimented upon much.

8.19 APPLICATIONS OF SUPERCONDUCTORS

Superconductors, particularly room-temperature superconductors, whenever discovered, will possess tremendous application potential. The vast range of applications of superconductors have attracted the attention of the researchers throughout the world. Room temperature superconductors will revolutionize the life and it will be one of the most important discovery (if discovered) in the history of mankind. While the room temperature superconductor will benefit almost all aspects of human life, even the presently available superconductors are increasingly being used despite the fact that they are required to be kept below their T_c and so involves expenditure in cooling them. The HTSC superconductors, with T_c greater than 77 K (liquid nitrogen temperature), have opened new opportunities for applications, as liquid nitrogen is much cheaper than liquid helium.

Although there are countless applications of superconductivity, a representative list is provided in Table 8.6. We discuss below a few important applications of superconductors.

Table 8.6 Important applications of superconductivity

Magnets	•	High-field magnet applications
	•	Nuclear magnetic resonance (NMR), medical diagnostics and spectroscopy
	•	Ore refining (magnetic separators)
	•	Magnetic levitation
	•	Magnetic shielding
	•	Large physics machines (colliders, fusion confinement, r.f. cavities)
Energy-related	•	Production by magnetic fusion and magneto-hydrodynamics
	•	Energy storage (magnetic)
	•	Electrical power transmission
Transportation	•	High-speed trains (Maglev)
	•	Ship-drive systems
Electronics and small devices	•	SQUIDS (Superconducting quantum interference devices)
	•	Josephson devices (square-law detector, amplifier, modulator, thermal valve)
	•	Bolometer
	•	Electromagnetic shielding
Computers and information processing	•	Semiconductor-superconductor hybrids (A-D converters)
	•	Active superconducting elements (FETs)
	•	Voltage standard
	•	Optoelectronics
	•	Matched filters

1. **Superconducting Magnets:** Magnets are one of the most useful applications of superconductors. Once the magnetic field has been set up on further electrical power is required to maintain it. For high magnetic field (> 10 Tesla) superconducting magnet is considerably more economical than conventional magnets. Superconducting magnets are used in nuclear magnetic resonance (NMR) spectrometers, magnetic resonance imaging (MRI) magneto-hydrodynamics (MHD), controlled nuclear fusion, elementary particle research, magnetically levitated (maglev) rail transportation, magnetic ship propulsion, magnetic separation, etc.

In maglev rail transportation, magnetic repulsion is used to lift and hold a train a little above the track and such a train actually flies over the track instead of rolling like in conventional trains. Speed upto 550 km/hour achieved in maglev trains. However, presently the expenditure on cooling, makes such trains economically unviable.

2. **Loss-less Power Transmission and Energy Storage:** Despite environmental concerns, there are big thermal power plants situated inside the big cities. Superconductors, particularly if available at room temperature, make it possible that power can be produced in the remote areas where coal/gas is available and the produced power can be transported to major consumption centre without any loss. Superconductors, being resistance free, do not dissipate any power. Efforts are being made to fabricate power cables and wires from the

available superconductors with high current carrying capacity. Superconducting electricity transmission use a superconductor inside an insulated pipe filled with liquid nitrogen.

Superconductors are also used for energy storage. If a current is introduced in a ring-shaped superconductor, the current will flow forever as long as superconductivity is maintained. Such storage rings can help reduce the fluctuations in the network. Such storage rings are used in particle accelerators such as cyclotron.

3. **Superconductors in Electronics and Computers:** The use of superconducting materials as coatings in radiowave generators, resonators and waveguides extends their frequency range. The use of HTSCs in semiconductors can increase the computing speed and help reduce the size of components. Use of superconductors can help reduce the heat generation in integrated chips (ICs).

Cryoton is used as (superconducting) switching device in computers. In this device, the switching is produced by magnetic suppression of superconductivity, which is much faster than that by using semiconductor transistors. Small cryotons can be used as fast acting switches in digital computers while large cryotons can be used to control currents in superconducting magnetic circuits. Superconductors can also be used as ultra-precise voltage standards, low power linear A/D converter, etc.

4. **Applications in Measurement Technology, Motors and Generators:** Superconducting galvanometer with very high sensitivity ($\approx 10^{-11}$ V), bolometers (radiation meters), amplifiers, modulators, thermal valve, etc. can be constructed using superconductors. Highly efficient motors and generators can be constructed using superconductors. Motors and generators, using superconductors, are expected to be of half the weight and half to one-third the size of conventional motors and generators in addition to being quieter and having lesser operating cost.

5. **SQUID:** A SQUID can act as ultra sensitive magnetometer capable of detecting even tiny magnetic change due the flow of current pulse owing to activity in brain. This has been used by the medical professionals for a variety of applications, e.g., monitoring the brain activity, detecting iron level in the lever, etc. It has been used for the detection of minerals and ores, oil exploration, etc. It has potential for being used in earthquake research and prediction and submarine detection.

At present, economic considerations are restricting large-scale commercial use of superconductors.

SHORT ANSWER TYPE QUESTIONS

1. What is *superconductivity*? Define critical transition temperature. What is its value for mercury?
2. Can we say superconductors have zero resistance? Explain.
3. What do you mean by *persistent current*? Explain.
4. Define *critical magnetic field* and discuss its variation with temperature.
5. Define *critical current* and *critical current density*.
6. What is *Meissner effect*? Explain.
7. *Meissner effect* and *non zero resistivity* is considered to be the sure signature of superconductivity. Why?

8. Only the superconductors are the perfect diamagnetic materials. Defend the statement.
9. Differentiate between type-I and type-II superconductors.
10. What do you mean by *penetration depth*? Explain.
11. Show the variation of the entropy of a superconductor. What important conclusion can be drawn from it.
12. The variation of specific heat with temperature is indicative of what? Elaborate.
13. Show the variation of the thermal conductivity with temperature for a material in superconducting and normal phase. How is this useful in understanding the phenomenon of superconductivity?
14. In what respect the energy gap in a superconductor differ from that in a normal metal? Explain.
15. What is *isotope effect*? Explain the significance of this effect.
16. Name the factors that determine the value of London penetration depth.
17. Mention three important features of the BCS theory.
18. What is fluxon? Explain.
19. What is SQUID? How it is useful?
20. Briefly distinguish between d.c. and a.c. Josephson effects.
21. Name the first high- T_c superconductor. Give its composition, T_c and characteristics.
22. How the discovery of superconductivity in 123 system was significant? Explain.
23. Give the compositions and T_c of Bi-, Tl- and Hg-based superconducting systems. What is the maximum T_c obtained so far?
24. What are superconducting fullerenes? Give the compositions of three fullerene and their T_c .
25. Mention the important applications of superconductors.

LONG ANSWER TYPE QUESTIONS

1. What do you mean by *superconductivity* and *superconductors*? Explain the discovery of superconductivity by K. Onnes.
2. Explain why a superconductor is a superconductor for direct current only and not alternating current.
3. 'Explain Meissner effect. How this effect is explained?
4. Distinguish between type-I and type-II superconductors giving suitable diagrams. What are ideal and non-ideal superconductors?
5. What is penetration depth? Discuss its variation with temperature.
6. Discuss the variation of specific heat of a superconductor. What important indication is obtained from this variation?
7. Discuss the important thermal properties of a superconductor and mention the important clues provided by them for understanding the superconductivity phenomenon.
8. Discuss the evidences suggesting the existence of an energy gap in a superconductor. How does this energy gap differ from that in a normal metal? Explain its variation with temperature.
9. Discuss the absorption of electromagnetic radiation by a superconductor and explain the observed behaviour.
10. Obtain London equations. Use these equations to get an expression for London penetration depth. Discuss the temperature dependence of the penetration depth.
11. Discuss the basic features of the BCS theory of superconductivity.

12. Explain the tunnelling between two superconductors separated by an insulating barrier. Hence, explain d.c. and a.c. Josephson effect.
13. What are high-temperature superconductors? Name different cuprate high- T_c superconductors and give their characteristics.
14. Why since the discovery of superconductivity, it has fascinated scientists and technologists? Explain in the context of wide ranging applications of superconductors.

EXERCISES

1. The critical transition temperature of a superconductor in the absence of magnetic field is T_c . At what temperature its critical magnetic field will reduce to its value at 0 K?
[Ans. 0.24 T_c]
2. The transition temperature of mercury is 4.185 K if its isotopic mass is 199.5. Estimate the T_c of mercury isotope with isotopic mass 203.4.
[Ans. 4.139 K]
3. Find the critical current density of a superconducting wire of radius 5 mm if its critical magnetic field is 8.5×10^3 A/m.
[Ans. 3.4×10^6 A]
4. Find the critical current density of a lead wire of 1 mm diameter at (a) 4.2 K and (b) 7 K. Given T_c of Pb is 7.18 K and its H_o is 6.5×10^4 A/m.
[Ans. (a) 1.7×10^8 A/m² (b) 1.36×10^7 A/m²]
5. The transition temperature of tin is 3.7 K in the absence of any magnetic field while its H_{co} is 0.0306 tesla. Estimate its critical magnetic field at 2 K.
[Ans. 0.05 A/m]
6. The transition temperature of Pb specimen is 7.22 K while its density and atomic weight is 11.3×10^3 kg/m³ and 207.19 K respectively. Find its London penetration depth at 0 K. Also calculate its value at 3.61 K and find the increase compared to its value at 0 K.
[Ans. 29.3 nm, 30.3 nm, 3.4%]
7. When a current I passes through a cylindrical superconducting wire of radius r , the magnetic field produced immediately outside the wire is H_c . If the current is measured in ampere, radius in centimetre and the field is Gauss show that $I = Sr H_c$.

9

Nanomaterials

9.1 INTRODUCTION

We are living in a ‘nano’ age where everything nano is considered to be exciting. Time is approaching to make a transition from ‘micro’ (e.g. microelectronics) to ‘nano’ fabrication age. In this chapter, we will understand why this field is so exciting, full of technological potentials.

Nano stands for 10^{-9} and by nano normally we mean one nanometre (10^{-9} m). *Materials with at least one of its dimensions measuring less than 100 nm (1 to 100 nm) are classified as nanomaterials.* If one dimension of a material is in nano range and the other two dimensions are large, then the structure of the material is referred to as *quantum well*. If the two of the three dimensions are in the nano range while the third is large, then the structure is referred to as *quantum wire*. When all the three dimensions are in nano range, the structure is referred to as *quantum dot*. Obviously, when none of the three dimensions is in nano range, i.e., all the dimensions are large the structure is a bulk structure.

Properties of the solids are dependent on their size. The properties of the bulk materials are mostly retained till the reduction of their dimensions in micrometre range. However, as the size of materials is reduced to nanometre range, the properties show remarkable change. Often surprisingly new properties are shown by the nanomaterials. For example, nano gold particles are not characteristic yellow but show different colours, orange, red, purple or greenish depending upon the actual size of the particles. Compared to bulk, the melting point and chemical properties also change when the material is reduced to nano form. The melting temperature in nano form reduces significantly. Bulk semiconductors become insulators when dimensions are reduced to nanometres. The primary cause for the drastic change in behaviour in nanomaterials compared to their bulk form is that in nano form the number of atoms on the surface is a big fraction of the total number of atoms in the material.

There is probably no sphere of human activity where the impact of the advancement of nano science and technology will not be felt. The advancement will impact all walk of human life. Particularly, as this field progress we will be moving from microelectronics age to nanoelectronics age. This has led to tremendous excitement and curiosity among the scientific community as well as general people towards the development in this field.

9.2 NANO SCIENCE & TECHNOLOGY: PAST AND FUTURE

Very often we develop perception that this emerging field of nanotechnology is new. For us this may be relatively new, but in fact the nature has its own nanotechnology and some of the nano-structures are old as nature itself. Many biological materials are nanoparticles. For example, viruses are in the nanometre range (10-200 nm) while the protein molecules lie in the range of 4-50 nm. Genetic material DNA (deoxyribonucleic acid) exists as compacted nanowire. The cells near the nose of trout fish are believed to possess a fluid of magnetic nanoparticles, which help these fish in finding direction. The fish takes advantage of the feel of change in magnetization when it deviates from the earth's magnetic field. Also the outer cover of the body of certain shells is extremely hard. It is believed it is because of the presence of a nano layer on the surface of the body.

Humans also started using the properties of nano size materials, much before the advent of modern science and technology. In fourth century A.D., Roman glass-makers fabricated the Lycrus cup, named after a king and now displayed in British museum in London. The cup contains nanosized metal particles and changes its colour from green to deep red on shining its inner surface with light. Thus, the field of nanomaterials is not new and the nanomaterials have been used by the nature and human before the advent of present form of nanoscience and technology.

In the modern documented science, the first use of the concepts of nanotechnology was in the famous talk, “*There is plenty of room at the bottom*” by Richard Feynmann at American Physical Society meeting at Caltech on December 29, 1959, wherein he talked the possibility of manipulating individual atoms and molecules. He further stated that a cube of side 0.02 inch could store all the books ever written if the material of the cube were constructed by adding individual atoms with the use of finely tuned technique. However, at that time, due to his nature of cracking jokes, he was not taken seriously by the audience. But he had rightly noted that scaling issues would arise from the changing magnitude of various physical phenomenon: gravity would become less important, surface tension and van der Waals attraction would become more important. The term nanotechnology was defined by Tokyo Science University Professor N. Taniguchi in 1974 as follows. According to Taniguchi, *Nanotechnology consists of the processing, separation, consolidation and deformation of materials by one atom or by one molecule*.

The size of the atoms are nearly 0.1 nm. The smallest naturally occurring atom, helium, is 0.1 nm while the size of uranium atom is 0.22 nm. In nanotechnology, one needs tools to see, measure and manipulate the atoms at the nanoscale. Some of the first tools to help launch the nanoscience revolution were the scanning probe instruments.

Some of the initial tools which helped nanoscience revolution were the scanning probe instruments, most of which conceived at the IBM Laboratory at Zurich. In scanning probe measurements, the probe (often called the tip) slides along the surface of the specimen. In scanning tunnelling microscopy (STM) the amount of electrical current between the scanning tip and the surface is measured. Depending on the way the measurement is performed, STM can be used either to know the local geometry (e.g., how much surface protudes locally) or to know the local electrical characteristics. STM was the first scanning probe method developed by Gerd Binning and Heinrich Rohrer for which they were awarded Nobel Prize. Scanning probes are used to manipulate individual atoms and molecules by either pushing them or picking them off the surface. In 1990, Donald Eigler of IBM and E.K. Schweizer from Fritz manipulated 35 xenon atoms to spell out three letter IBM logo on a nickel crystal (Fig. 9.1).

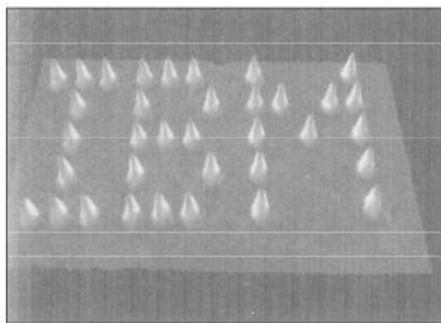


Fig. 9.1 Positioning single atoms with a scanning tunnelling microscope.

From vacuum tubes to transistors to integrated circuits (IC) to very large-scale integration (VLSI), there have been a steady decline in the size of electronic devices (miniaturization of the electronic devices) accompanied by enhanced performance. If the trend is extrapolated, the days are not distant when the number of atoms required to write one bit in a mass memory device reaches one, while the energy required for a single logic operation reaches kT (for $T = 300$ K) which is the energy possessed by a single molecule at room temperature.

The finest machine technology will thus reach a resolution of an atomic diameter.

Gordon E. Moore, one of the founders of Intel Corporation in an article titled “Cramming more components onto integrated circuits” gave two laws, known as *Moore’s laws*. According to these laws:

1. The amount of space required to install a transistor on a chip shrinks roughly half every 18 months. This means that the space required for one transistor 15 years ago, can hold 1000 transistors today.
2. The cost of fabricating a chip doubles with every other chip generation or roughly every 3 years.

The Moore’s first law has largely been true. From just over 2000 transistors in the original 4004 processors of 1971 to over 40,000,000 transistors in the Pentium 4.

Moore had also realized that we are approaching some real limits. Further miniaturization will be possible only with nano scale electronics or molecular electronics. Molecular electronics can be defined as a subfield of nanotechnology that envisions the use of single molecules, or small group of molecules, as components in electronics applications. Molecular electronic devices would form the next generation of transistors, sensors and circuits. If this would be possible, the validity of Moore’s law could be extended.

Nanotechnology holds promise of tremendous applications such as targeted drug delivery, self cleaning buildings, highly efficient nanofilters, nanoswitches, nanocomputers, nano electromechanical systems (NEMS), etc.

9.3 NANOPARTICLES: DEVIATION FROM BULK BEHAVIOUR

Nanoparticles are generally considered to be the structures, consisting of atoms or molecules, with a radius less than 100 nm (or 1000 \AA as $1 \text{ nm} = 10 \text{ \AA}$). In Fig. 9.2 a general classification of atomic

cluster, according to the number of atoms present in the cluster, is given. When the number of atoms is less, most of them are on the surface. For example, a cluster comprising 13 atoms, in FCC structure, 92% of the atoms are on the surface. As the number of atoms increases in a cluster, the percentage of atoms on the surface decreases. *It is the presence of large percentage of atoms on the surface in the nanostructures, which make them behave differently compared to the bulk structures.*

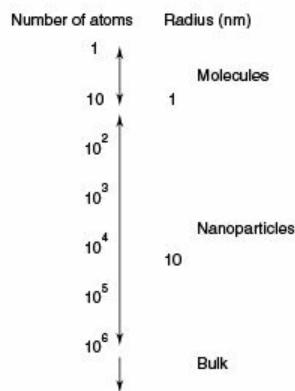


Fig. 9.2 Classification between molecules, nanoparticles and bulk on the basis of the number of atoms in the cluster.

Also, the physics and chemistry of nanostructures are different than that of the bulk structures. Particles of different sizes scatter different wavelengths of light. This fact has been used from ancient times to produce beautiful colours in stained-glass, due to the presence of the nano metal-clusters in the glass having size comparable to the wavelength of light. What makes nanoparticles very interesting and endows them with their unique properties is that their size is smaller than critical lengths that characterize many physical phenomena. Generally, physical properties are associated with some characteristic length such as scattering length, thermal diffusion length, etc. The electrical resistance of a metal is dependent upon the mean free path or scattering length, the electrons travel between any two successive collisions. If the sizes of the particles are less than these characteristic lengths, the physics of the process changes.

9.4 METAL NANOCLUSTERS

In Fig. 9.3, the process for the formation of metal nano particles or clusters are depicted on a metal rod high intensity laser beam is incident which causes the evaporation of atoms from the surface. A jet of helium gas is used to force the evaporated atoms into an evacuated chamber where the expansion of the gas causes cooling and the metal atoms form clusters. The clusters are then ionized by UV radiation and then passed into a mass spectrometer that measures the charge to mass ratio using which the number of atoms in the clusters and the stability of the cluster is determined. The clusters can have 2 to 1000 of atoms.

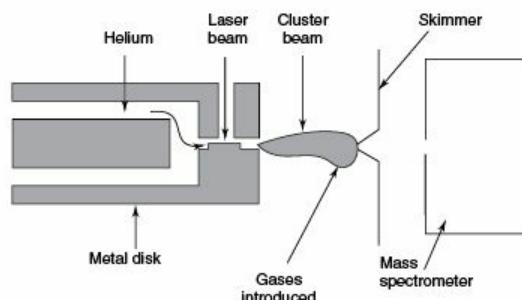


Fig. 9.3 Formation of metal nanoparticles by laser induced evaporation of atoms from metal surface.

It is observed that there are large number of clusters having certain specific number of atoms. These numbers indicate that certain specific number of atoms form a stable cluster. These numbers are called *magic numbers* or *structural magic numbers*.

Since the clusters are small aggregates of atoms, the number of atoms on the surface are higher than the number of atoms inside, i.e., the surface to volume ratio in these aggregates of atoms is quite large compared to the bulk. For a cluster of 1000 atoms, 25% atoms lie on the surface.

There are two well-known theoretical models to calculate the properties of clusters: (i) jellium model (or approximation) and (ii) density functional model.

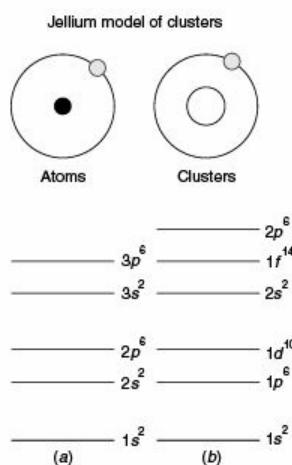


Fig. 9.4 (a) Energy levels of hydrogen atom and (b) Energy levels of a cluster according to jellium model.

The Jellium model: In this model a cluster of atoms is treated as a large atom, often called a *superatom*. The positive nuclear charge of each atom of the cluster is assumed to be uniformly distributed over a sphere of the size of the cluster. The interaction between the electron and the positive spherical charge distribution is represented by a spherically symmetric potential well. Therefore, the energy levels can be obtained by solving the Schrödinger equation in a manner similar to that for the hydrogen atom. However, the solution of Schrödinger equation yields a different order of energy levels for clusters, as depicted in Fig. 9.4. Instead of $1s^2, 2s^2, 2p^6, 3s^2$ and $3p^6$ for the hydrogen atom. The energy levels for a cluster is found to be $1s^2, 1p^6, 1d^{10}, 2s^2, 1f^{14}$ and $2p^6$. While the electronic magic number for the clusters are found to be 2, 18 and 40 instead of 2, 10, 18 and 36 for He, Ne, Ar and Kr respectively. In the Jellium model, the magic numbers correspond to those clusters having a size in which all the energy levels are filled. These magic numbers are also referred to as *structural magic numbers*.

Molecular orbital theory approach: An alternative approach to calculate the properties of clusters is based on the principles of molecular orbital theory. The simplest cluster possible is a H_2^+ molecule. This theory assumes that the wavefunction of the electron around the two H nuclei can be represented by a linear combination of the wavefunction of the isolated H atoms. Thus, the wavefunction in the ground can be represented as

$$\Psi = \alpha \psi(1)_{1s} + \beta \psi(2)_{1s} \quad \dots(9.1)$$

The Schrödinger equation of the molecular ion is

$$\left[\left(-\frac{\hbar^2}{2m} \right) \nabla^2 - \frac{e^2}{r_a} - \frac{e^2}{r_b} \right] \Psi = E \Psi \quad \dots(9.2)$$

The symbol $\nabla^2 \left(\equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$ denotes a double differentiation operation. The last two terms in the brackets are the electrostatic attraction of the electron to the positive nuclei which are at distances r_a and r_b from the electron. In case of a hydrogen molecule, which consists of two electrons, another term for the electrostatic repulsion of the two electrons is added. The Schrödinger equation is solved with the linear combination of wavefunctions (Eq. 9.1). For a molecule consisting of many electrons, the problem becomes complex and many approximation has to be used to obtain the solution.

9.5 QUANTUM STRUCTURES: DIMENSION EFFECTS

When the size of a material is reduced to nanometre range (< 100 nm), there is a marked change in the properties of the material. If one dimension is reduced to the nanorange but the other two dimensions remain large, then the structure is known as a *quantum well*. If two dimensions are reduced to nanorange but one of the dimensions remains large, the resulting structure is *quantum wire*. When all the three dimensions are in nanometer range, the structure is referred to as *quantum dot*.

It is the large percentage of atoms on the surface in the nanostructures compared to the bulk structures which differentiates properties of nanostructures compared to the bulk material. Figure 9.5 shows that the percentage of surface atoms increases sharply as the particle size reduces to less than 5 nm, for the spherical iron nanoparticles.

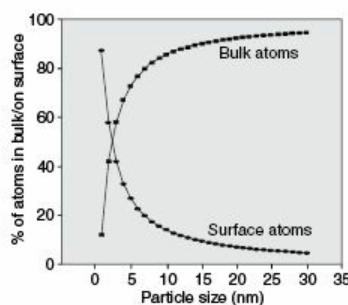


Fig. 9.5 Spherical iron nanocrystals—particle size vs % atoms in bulk/on surface.

In large systems, conduction or free electrons is delocalised as all the wire dimensions are very large compared to the distances between atoms. As size reduces, electrons get confined and delocalisation decreases. A copper plate of 10 cm long, 10 cm wide and 3.6 nm thick contains 10 unit cells and only 20% of the atoms in unit cells are at the surface. In such a plate, the conduction electrons would be delocalised in the plane of the plate but confined in the narrow (thickness) dimension and this configuration is referred to as **quantum well**. A quantum wire is a structure that is long in one dimension but its diameter is in nm. The electrons are delocalized and move freely along the wire but are confined in the transverse directions. A quantum dot exhibits electron confinement in all directions.

We discuss here the confinement of a particle in one dimension (other two dimensions large-particles free to move), two dimensions (third dimension large-particle free to move in this direction) and three dimensions (no free movement in any direction as in OD material). In quantum mechanics, we deal with such cases as particle in a potential box. These confinement cases are also known as 2D quantum well (thin film), 1D (tube or wire) and OD quantum dot (nanoparticles).

It is relevant to recall that for a box of length a with potential $V=0$ for $0 < x < a$ and $V=\infty$ for $x < 0$ or $x > a$ (one-dimensional potential box), the solution of the Schrödinger equation yields eigenvalues

$$E_n = \frac{n^2 \hbar^2 \pi^2}{2ma^2} = \frac{n^2 \hbar^2}{8ma^2} \quad \text{or} \quad E_n \propto n^2$$

and eigenfunctions

$$\psi_n = A \sin\left(\frac{n\pi}{a}\right)x \quad (n = 1, 2, 3, \dots)$$

The first three energy levels and corresponding wavefunctions and probabilities are shown in Fig. 9.6 respectively. For nanoparticles, the particle behaves like an individual atom possessing discrete energy levels rather than continuum.

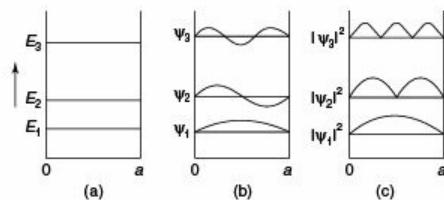


Fig. 9.6 (a) Quantized energy states (b) Corresponding wavefunctions and (c) Probability of finding the particle at different locations between '0' and 'a' in the box.

9.5.1 Density of States

Density of states $D(E)$ is defined as the number of states per unit energy range (E and $E + dE$). It is an important parameter as it helps in understanding various spectroscopic and transport properties of materials. As the energy of a particle in a one-dimensional box is

$$E_n = \frac{\hbar^2}{8ma^2} n^2$$

or $dE = \frac{\hbar^2}{8ma^2} 2n dn$ (ignoring suffix n with E)

$$\therefore \frac{dn}{dE} = \frac{8ma^2}{\hbar^2} \frac{1}{2n} = \frac{a}{\hbar} \sqrt{\frac{2m}{E}}$$

or $D(E) \left(\equiv \frac{dn}{dE} \right) \propto E^{-1/2}$... (9.3)

This shows that the density of states vary with energy and it is inversely proportional to the square root of energy. We discuss below the variation of density of states for solids of different dimensions.

9.5.2 Density of States for a Zero Dimensional (0D) Structure

In zero dimensional structures, like quantum dots, an electron is confined in a three-dimensional potential box with extremely small (< 100 nm) length, breadth and height as a 0D solid. All three spatial directions are quantized and the expression for the density of states follows a δ -function given as

$$D(E) = \sum_i \delta(E - \varepsilon_i)$$

where the summation is over all the i quantum states and ε_i are the discrete energy levels and δ is Dirac function.

The density of state as a function of energy appears as shown in Fig. 9.7.

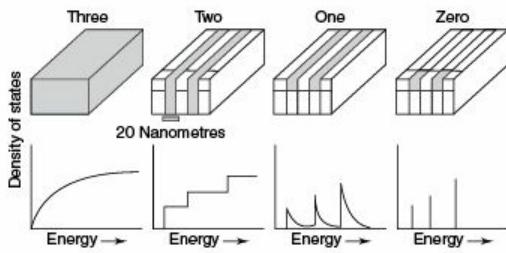


Fig. 9.7 Quantum structures—density of states vs energy.

9.5.3 Density of States for a One-Dimensional (1D) Structure

For 1D structures like quantum wire, a particle (e.g., charge carrier) can move in one direction and in other two directions it is confined. The potential in two directions is infinitely large but in the third direction it is zero. The density of states in this case is expressed as

$$D(E) = \frac{dN}{dE} = \sum_{\epsilon_i < E} (E - \epsilon_i)^{1/2} \quad \dots(9.4)$$

where ϵ_i are discrete energy levels. The graphical representation of density of states for 1D structure is shown in Fig. 9.7.

9.5.4 Density of States for Two-Dimensional (2D) Structures

For 2D structures which is basically a thin film structure (two directions long, one direction < 100 nm). The density of states is given by

$$\begin{aligned} D(E) &= 0 \text{ for } E < \epsilon_i \\ &= 1 \text{ for } E > \epsilon_i \end{aligned}$$

where ϵ_i is the i th energy level within 2D quantum well. The graphical representation of density of states for 2D structure, a staircase structure, is shown in Fig. 9.7.

9.5.5 Density of States for Three-Dimensional (3D) Structures

For a 3D structure which is equivalent to a three-dimensional box of length a , width b and height c with potential $V = 0$ inside the box and $V = \infty$ outside the box, the solution of Schrödinger equation shows that the energy states are:

$$E_{nx, ny, nz} = \frac{\hbar^2}{2m} (n_x^2 + n_y^2 + n_z^2)$$

while the corresponding wavefunction has the form

$$\psi_n(x, y, z) = A \sin\left(\frac{\pi n_x x}{a}\right) \sin\left(\frac{\pi n_y y}{b}\right) \sin\left(\frac{\pi n_z z}{c}\right) \quad \dots(9.5)$$

It can be shown that the density of states in this case

$$D(E) \propto E^{1/2}$$

which is represented graphically in Fig. 9.7.

9.6 CARBON NANOSTRUCTURES

9.6.1 Carbon Molecules

Carbon atoms have the ability to form different structures. Carbon forms many nanostructure, carbon nanotubes being the most important because of their useful and interesting properties. The two main structures of solid carbon, called allotropic forms, which are stable at room temperature, are *diamond* and *graphite*. Diamond consists of carbon atoms that are tetrahedrally bonded to each other through sp^3 hybrid bonds (bond angle 109°) forming a three-dimensional network. Thus, each carbon atom has four nearest-neighbour carbons. Graphite has a layered structure, with each layer called a *graphite sheet*, formed from hexagons of carbon atoms bound together by sp^2 hybrid bonds that make 120° angles to each other. Each carbon atom has 3 nearest neighbour carbon atoms in the planar layer (Fig. 9.8). The hexagonal sheets are held together by weak van der Waals forces.

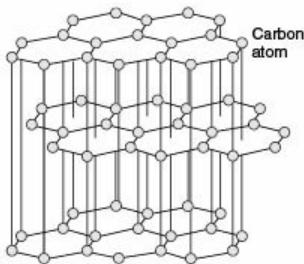


Fig. 9.8 Structure of graphite.

For fairly long time, it was believed that no other carbon bond angles (other than 180° for linear acetylene $H - C \equiv C - H$, 120° for planar ethylene $H_2C = CH_2$ and 109° 28' for the tetrahedral methane CH_4 , respectively). However, in 1964, Phil Eaton of the University of Chicago synthesized a square carbon molecule, C_8H_8 , called *cubane*. In 1983, L. Paquette of Ohio State University synthesized a $C_{20}H_{20}$ molecule having a dodecahedron shape formed by joining carbon pentagons and having C – C bond angles ranging from 108° to 110°. The synthesis of these new hydrocarbon molecules with carbon bond angles different from the standard hybridization values opened the possibilities of synthesizing new molecules leading to the synthesis of carbon nanotubes.

9.6.2 Discovery of C_{60} (Fullerenes)

Theoretical chemists had envisioned the existence of a soccer ball-like molecule consisting of 60 carbon atoms with the chemical formula C_{60} . Some of them had even predicted the properties of this molecule. The actual existence of this molecule was reported in 1990 in an article in the *Nature*. Later an experiment was carried out to synthesize C_{60} molecule by Prof. Richard Smalley of Rice University in collaboration with Harold Kroto, a chemist from the University of Sussex. In this experiment, a graphite disk is heated by a high intensity laser beam which produces a hot vapour of carbon. A burst of helium gas then sweeps the vapour out through an opening where the beam expands. The expansion cools the atoms and they condense into clusters. This cooled cluster beam

is then narrowed by a skimmer and fed into a mass spectrometer, which is an instrument used to measure the mass of molecules in the clusters. A mass number of 720 consisting of 60 carbon atoms, each of mass 12, was observed, confirming the existence of C_{60} molecule. The C_{60} molecule was named after an architect R. Buckminster Fuller, who had designed dome that resembles the structure of C_{60} . Thus, the molecule C_{60} was originally called *buckminsterfullerene*, but it being too long, very often the term *fullerene* is used.

Each C_{60} molecule is composed of a group of carbon atoms that are bonded to one another to form both hexagon (six-carbon atom) and pentagon (five-carbon atom) geometrical configurations. One such molecule, shown in Fig. 9.9, is found to consist of 20 hexagons and 12 pentagons, which are arrayed such that no two pentagons share a common side. Thus, the molecular surface exhibits the symmetry of a soccer ball, often also referred to as **bucky ball**.

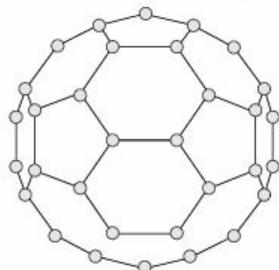


Fig. 9.9 Structure of a C_{60} fullerene molecule.

These ball-like molecules bind each other in the solid state to form a crystal lattice having a FCC structure. In contrast to diamond and graphite, which are networked solids in which all the carbon atoms form primary bonds with adjacent atoms throughout the solid, the carbon atoms in fullerene bond together to form spherical molecules. Each C_{60} molecule separated from its nearest neighbour by 1 nm are held together by weak van der Waals forces. C_{60} is soluble in benzene and single crystals can be grown by evaporation of benzene.

As a pure crystalline solid, fullerene is electrically insulating. However, with proper impurity additions, it can be made highly conductive (or superconducting at low temperatures) and semiconducting. Twenty-six per cent of the volume of the unit cell is empty so alkali atoms can easily fit into empty spaces between the molecular balls of the material. When C_{60} crystal and potassium metal are heated to 400°C in an evacuated chamber, potassium vapour diffuses into these empty spaces to form the compound $K_3 C_{60}$, and becomes electrically conducting. In $K_3 C_{60}$, the potassium atoms become ionized to form K^+ and their electrons get attached to C_{60} forming triply negative ion C_{60}^{3-} . Thus, each C_{60} has extra electrons that are loosely bonded to the C_{60} molecule (C_{60} called electron doped) and can move through the lattice making the material electrically conducting.

In 1991, A.F. Hebard and his co-workers at Bell Telephone Laboratories doped C_{60} crystals with potassium and found that the material shows superconductivity with a transition temperature of 18 K. Later it was found that many alkali atoms could be doped into the lattice and the transition temperature increased to 33 K in the compound $Cs_3 Rb C_{60}$. As the radius of the dopant alkali atom increases, the cubic C_{60} lattice expands, and the superconducting transition temperature increases.

Larger fullerenes such as C₇₀, C₇₆, C₈₀ and C₈₄ are also known. A C₂₀ dodecahedral carbon molecule has been synthesized by gas-phase dissociation of C₂₀HBr₁₃. C₃₆H₄ has been synthesized by pulsed laser ablation of graphite. A solid phase of C₂₂ has been identified in which the lattice consists of C₂₀ molecules bonded together by intermediate carbon atom. Some of these smaller fullerenes have the potential to be high-temperature superconductor when doped appropriately.

9.6.3 Carbon Nanotubes (CNTs)

Carbon nanotubes are nanostructures with large application potential discovered in 1991 by Sumio Iijima of NEC laboratory in Japan. Its structure consists of a single sheet of graphite, rolled into a tube (cylinder), both ends (or at least one) of which are capped with C₆₀ fullerenes hemispheres (Fig. 9.10).

A typical *single-walled nanotube* (SWNT) can have a diameter of 2 nm and a length upto 100 μm, and effectively it is a one-dimensional structure called a *nanowire*. Sometimes many coaxial concentric cylinders may exist as carbon nanotubes. Such concentric nanotubes are termed Multi Wall Carbon Nanotubes (MWCNTs). The separation between the layers is around 0.334 nm, which is nearly the same as the separation between two graphite layers in a single crystal. In the synthesis process most of the times MWCNTs are formed, however, under certain conditions it is possible to obtain single wall carbon nanotubes (SWCNTs). The diameter of multiwalled tubes can extend upto 25 nm.

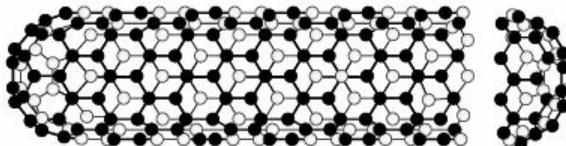


Fig. 9.10 Structure of a carbon nanotube.

Carbon nanotubes can exist in different structures and these structures yield different properties. Although carbon nanotubes are not actually made by rolling graphite sheets, however, it is possible to explain different structures by considering the way graphite sheets can be rolled into tubes. A nanotube can be formed by rolling a graphite sheet about the axis T (called axis vector T), shown in Fig. 9.11. The vector C_h in Fig. 9.11 is called the *circumferential vector*, which is at right angle to the vector T. Three examples of nanotube structures formed by rolling the graphite sheet about the T vector different orientations in the graphite sheet are shown in Fig. 9.12. When T is parallel to the C–C bonds of the carbon hexagons, the structure shown in Fig. 9.12(a) is obtained and it is referred to as the *armchair* structure. The structures shown in Fig. 9.12(b) and (c) are

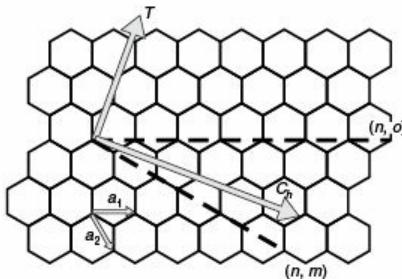


Fig. 9.11 Graphite sheet with basis vectors a_1 and a_2 and circumferential vector C_h .

respectively referred to as zigzag (along the horizontal zigzag bonds; along the dotted (n, o) and chiral (along the dotted line (n, m) nanotubes. In terms of the values of n and m , the armchair nanotube is formed when $n = m$. The zigzag nanotube forms when $m = 0$. While the chiral type nanotube occurs when the numbers (n, m) can assume any integer values and the chiral angle (the angle between circumferential vector and the unit vector a) is intermediate between 0 and 30. These are obtained by rolling about a T vector having different orientations in the graphite plane not parallel to C – C bonds. Looking down the tube, in the chiral structure spiralling row of carbon atoms are seen. The tubes are generally closed at both ends with each end of the cylindrical tube attached to half of a large fullerene structure. Among the three structures, armchair SWCNTs are metals, those with $n - m = k$ (k is an integer) are semiconductors with a small band gap and the remainder are semiconductors with a band gap that is inversely proportional to the CNT diameter.

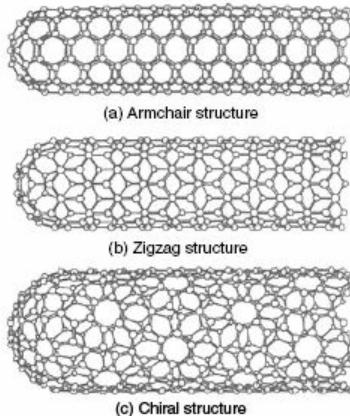


Fig. 9.12 Structures of carbon nanotubes.

Table 9.1 Comparison of synthesis routes of CNTs

Method	Type of CNTs	Diameter	Length	Advantages	Disadvantages
Laser	SWNT	1-2 nm	Long	Good size control Less defects	Very expensive
Vaporization					
Arc discharge	SWNT/MWNT	0.6-1.4 nm/10 nm	Short	Easy method Less defects	Short length Random size
CVD	SWNT/MWNT	0.6-4 nm/10-240 nm	Long	Easy to produce	Large defects Usually MWNT

9.6.4 Synthesis of CNTs

The most common methods used for synthesis of CNTs are: (i) laser evaporation, (ii) carbon arc discharge and (iii) chemical vapour deposition (CVD). Table 9.1 gives comparative advantages and

disadvantages of CNTs synthesized using these methods. The synthesis techniques are discussed below.

Figure 9.13 shows a schematic experimental arrangement for synthesizing carbon nanotubes by laser evaporation. A quartz tube containing argon gas and a graphite target are heated to 1200°C. In the tube, beyond the heated region of the furnace, there is a water-cooled copper collector. A small amount of cobalt and nickel is contained in the graphite target which acts as catalytic nucleation sites for the formation of the tubes. A highly intense beam of laser is incident on the target, evaporating carbon from graphite. The argon then sweeps the carbon atoms from the high-temperature zone to the colder copper collector on which they condense into nanotubes.

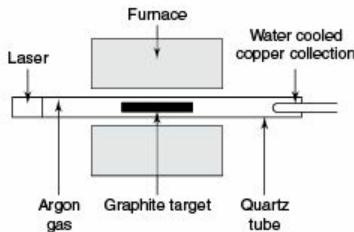


Fig. 9.13 Schematic diagram for synthesis of CNT by laser evaporation.

Nanotubes are also synthesized using carbon arc discharge. A potential of 20-25 V is applied across carbon electrodes of 5-20 µm diameter and separated by 1 mm at 500 torr pressure of flowing helium. Carbon atoms are ejected from the positive electrode and form nanotubes on the negative electrode. As the tubes form, the length of the positive electrode decreases and a carbon deposit forms on the negative electrode. To produce SWCNTs, a small amount of cobalt, nickel or iron is used as catalyst by putting it in the central region of the positive electrode. In absence of a catalyst, the tubes are nested or MWCNTs are formed.

The chemical vapour deposition method involves decomposing a hydrocarbon gas such as methane CH₄ at 100°C. As the gas decomposes, carbon atoms are produced which then condense on a cooler substrate that contains catalysts such as iron. This method produces tubes with open ends unlike other methods. This procedure allows continuous fabrication and it is suitable for large-scale production.

The mechanism of nanotube growth is not fully understood. Since the metal catalyst is necessary for the growth of SWCNTs, it is believed that metallic atoms such as CO or Ni are involved in the growth process. One hypothesis, referred to as the *scooter mechanism*, suggests that the atoms of the metal catalyst attach to the dangling bonds at the open end of the tubes and these atoms scoot around the rim of the tube absorbing carbon atoms as they arrive.

9.6.5 Properties of CNTs

Electrical Properties: Carbon nanotubes can be metallic or semiconducting, depending on the diameter and chirality (chirality refers to how the tubes are rolled with respect to the direction of the *T* vector in the graphene sheet) of the tube. Synthesis usually yields a mixture of tubes two-thirds of

which are semiconducting and one-third metallic. The metallic tubes have the armchair structure while the semiconducting tubes have chiral structure. The energy band gap in semiconducting chiral nanotube decreases with the increase of the tube diameter, varying from nearly 0.05 eV to 1.2 eV.

The dominant conduction path in CNTs is along the axis of the tubes, making carbon nanotubes function as one-dimensional quantum wires. The electronic states of the tubes do not form a single wide electronic energy band, but instead splits into one-dimensional subbands. Thus, these states can be modelled by a potential well having a depth equal to the length of the nanotube.

In the metallic state, the conductivity of the nanotubes is extremely high, of the order of 10^9 A cm^{-2} compared to 10^6 A cm^{-2} in copper (at higher current copper wire melts due to resistive heating). *The reason for the high conductivity of the carbon tubes is that they have very few defects to scatter electrons and therefore shows very low resistance.* High currents do not heat the tubes in the same way that they heat copper wires. Nanotubes have high thermal conductivity and so they are good conductors of heat.

Carbon nanotubes display *magnetoresistance* (a phenomenon in which the resistance of a material changes on the application of a magnetic field) at low temperature. In fact, it shows negative magnetoresistance as the resistance decreases with increasing magnetic field. In other words, its conductance increases on the application of a d.c. magnetic field. This is because when a d.c. magnetic field is applied to the nanotubes, the conduction electron acquires new energy levels associated with their spiralling motion about the field. For nanotubes these levels, called *Landau levels*, lie very close to topmost filled levels (the Fermi level). Therefore, there are more available states for the electrons to increase their energy and the material is more conducting.

Mechanical Properties: Carbon nanotubes are about 20 times stronger than steel. The tensile strength (a measure of the amount of stress needed to pull a material apart) of carbon nanotubes is about 45 billion pascals compared to about 2 billion pascals for high strength steel.

Young's modulus (= stress/strain) is a property of a material which characterizes the elastic flexibility of a material. The larger the value of Young's modulus, the less flexible the material. Young's modulus of steel is about 30000 times that of rubber. Compared to 0.21 terapascal (TPa) for steel the Young's modulus of carbon nanotubes range from 1.28 to 1.8 TPa (one terapascal is a pressure very close to 10^7 times atmospheric pressure). Thus, the Young's modulus of carbon nanotubes is almost 10 times that of steel. This would, however, imply that carbon nanotubes are very stiff and hard to bend. But this is not so as the carbon nanotubes are very thin.

When carbon nanotubes are bent, their behaviour is very resilient. They buckle like straws but do not break and can be straightened back without any damage. Most of the materials fracture on bending because of the presence of defects such as dislocations or grain boundaries. However, the structure of carbon nanotubes has so few defects, this does not occur. Another reason why they do not fracture is that as they are bent severely, the almost hexagonal carbon rings in the walls change in structure but do not break. This is due to the fact that the carbon-carbon bonds are sp^2 hybrids and these sp^2 bonds can rehybridize as they are bent. The degree of change and the amount of *s-p* admixture depend on the degree of bending of the bonds.

9.6.6 Application of Carbon Nanotubes

Due to outstanding and unusual properties carbon nanotubes are likely to be used for a variety of applications such as nanoelectrics, nanocomposites, field emission devices, nanosensors, hydrogen

storage, reinforcing fibres, battery electrodes, etc. However, the large-scale production of single walled carbon nanotubes is yet to be realized and presently it is very costly. We discuss below some of the potential applications of CNTs.

Field Emission and Shielding

When a small electric field is applied parallel to the axis of the tube, electrons are emitted at a very high rate from the ends of the tube. This is called *field emission*. This effect can be observed by applying a small voltage between two parallel metal electrodes and spreading a composite paste of nanotubes on one electrode. A sufficient number of tubes will be perpendicular to the electrode so that the electron emission can be observed. One application of this effect is the development of *flat panel displays*. Television and computer monitors use a controlled electron gun to impinge electrons on the phosphors of the screen which then emit visible light of appropriate colours. Samsung and Motorola is developing a flat-panel display using the electron emission of carbon nanotubes. A thin film of nanotubes is placed over control electronics with a phosphor coated glass plate on top. Notwithstanding certain problem, displays using CNTs are superior than conventional LCDs due to high brightness, a wide angle of view and low power consumption. A Japanese company is using this electron emission effect to make vacuum tube lamps that are as bright as conventional light bulbs but have longer life and more efficient.

Carbon nanotubes are poor transmitters of electromagnetic radiations due to their high electrical conductivity. A plastic composite of carbon nanotubes could provide lightweight shielding material for electromagnetic radiation, which is very important for applications.

In Nanoelectronics: CNTs are useful as the new generation *interconnect structures* as well as *field-effect transistors* (FETs). Interconnects which carry the electrical signals between transistors are currently made of copper, but as electronic circuits continue to shrink, copper interconnects will suffer from overheating. Conventional metal wires can typically exhibit current density of 10^5 A/cm^2 beyond which heating becomes a problem. On the other hand, because of the nearly one-dimensional electronic structure of CNTs, electronic transport in metallic SWNTs occurs ballistically along the nanotube length, allowing the conduction of high currents with no heating. In fact, current densities upto 10^9 A/cm^2 have been observed in SWCNTs. This is because the electronic states are confined in the directions perpendicular to the tube axis. Due to the lack of phonon and/or impurity scattering perpendicular to the tube, CNTs behave as 1-D ballistic conductors.

Field-effect transistors, used as switching components in computers, are normally made up of silicon and are of a few hundred nanometres in size. The use of CNTs with size less than 1 nanometre in diameter would allow more of these switches to be placed on a chip. In an FET, the current flows through a CNT with a semiconductor properties along a path called the *channel* (Fig. 9.14). At one side of the channel is a gold electrode called the *source* and at the other side of the channel is a gold electrode called the *drain*. When a small voltage is applied to the silicon substrate, which acts as a gate in FETs, the conductivity of the CNT can change more than a million times, which is comparable to silicon field-effect transistors. It has been estimated that the switching time of these devices will be very fast, allowing clock speeds of terahertz, which is 10^4 times faster than the present processors. The gold sources and drains are deposited by lithographic methods and the connecting nanotube wire is less than one nanometre in diameter. This small size makes it possible for more switches to be packed on a chip.

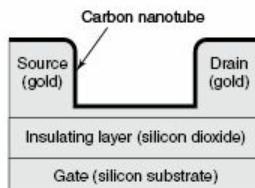


Fig. 9.14 A FET using CNT.

Also in the area of nanoelectronics, a challenging idea is to build entire electronic circuit out of CNTs, making use of their metallic and semiconducting properties. For this, the semiconducting CNTs are aligned on an insulator substrate, whereas metallic CNTs are placed above in the close proximity to the bottom layer. By controlling the current, the top CNTs can be made to contact the bottom CNTs, producing a metal semiconductor junction that acts as a switch. In this way, it is estimated that 10^{12} switches could fit on a square centimetre chip. The present Pentium chips have about 10^8 switches on them, indicating the possibility of further miniaturization. The switching rate of such devices is estimated to be about 100 times faster than that of the present generation of Intel chips.

In Fuel Cells: Carbon nanotubes are useful in fuel cells as well as in batteries, primarily for storage purposes of hydrogen. Fuel cells are being developed as sources of electrical energy for future automobiles. A fuel cell consists of two electrodes separated by a special electrolyte that allows hydrogen ions, but not electrons, to pass through it. Hydrogen when reaches anode gets ionized. The electrons released in the process move through an external circuit to the cathode. The hydrogen ions diffuse through the electrolyte to the cathode, where electrons, hydrogen and oxygen combine to form water. The system needs a source of hydrogen which can be stored inside carbon nanotubes. It is estimated that to be useful in this application, the tubes need to hold 6.5% hydrogen by weight. At present only about 4% hydrogen by weight has been successfully put inside the tubes.

Lithium which is charge carrier in some batteries can be stored inside the carbon nanotubes. The abundant open structure of CNTs is very useful for storage of large amounts of lithium ions. The properties of CNTs such as good chemical stability, large surface area and elastic modulus are helpful in prolonging the lifespan of the batteries based on CNTs.

Mechanical Reinforcement: Use of carbon nanotubes to increase the tensile strength of many materials has been established by many research groups. The researchers believe a substantial increase in the tensile strength can be achieved by a homogeneous and aligned distribution of nanotubes in the material. It is estimated that using 10% volume fraction of nanotubes the tensile strength can be increased upto 6 times.

Supercapacitors: CNTs can be of great use in the field of supercapacitors as CNTs exhibit high porosity, large specific surface area, high electrical conductivity and chemical stability. In a parallel plate capacitor, the capacitance depends on the separation between the plates and the dielectric material inserted between the plates. In the case of a supercapacitor, there is instead an electrical double layer (Fig. 9.19). Each layer contains a highly porous electrode suspended within an electrolyte. An applied potential on the positive electrode attracts the negative ions in the electrolyte, whereas the potential on the negative electrode attracts the positive ions. A dielectric material between the two electrodes prevents the charges from crossing between the two electrodes. If the electrodes are

made of CNTs, the effective charge separation is about a nanometre, compared with separations of the order of micrometres for ordinary capacitor. This small separation, combined with large surface area, is responsible for the high capacitance (upto two orders of magnitude higher than conventional capacitor) of these devices. In addition, although it is an electrochemical device, no chemical reactions are involved, allowing the ultra capacitor to be rapidly charged and discharged. Supercapacitors employing MWCNT electrodes have already achieved a capacitance ranging from 18 to 250 F/g.

Chemical Sensor: Field-effect transistor made of the chiral semiconducting CNTs has been found to be a sensitive detector of various gases. The frequency of one of the normal-mode vibrations of the nanotubes, which produces a very strong Raman line, is very sensitive to the presence of other molecules on the surface of the tubes. The direction and magnitude of the shift depend on the nature of the molecule on surface of the CNTs. This effect is the basis of the using nanotubes as chemical gas sensor.

Catalyst: Any material, normally a metal or an alloy, which enhances the rate of a reaction is called a catalyst. For some chemical reactions, CNTs can be used as catalyst. For example, nested nanotubes with ruthenium metal bonded to it, have been shown to have a strong catalytic effect in the hydrogenation reaction of cinnamaldehyde ($C_6H_5CH=CH CHO$) in the liquid phase compared to the effect when the same metal Ru is attached to other carbon substrates.

In fact, there are enormous number of applications and more are being explored continuously.

9.7 OTHER APPLICATIONS OF NANOMATERIALS

Next-Generation Computer Chips

The microelectronics industry has been emphasizing miniaturization, whereby the circuits, such as transistors, resistors, capacitors, etc. are reduced in size. By achieving a significant reduction in their size, the microprocessors, which contain these components, can run much faster, thereby enabling computations at far greater speeds. However, there are several technological impediments to these advancements, including lack of the ultrafine precursors to manufacture these components; poor dissipation of tremendous amount of heat generated by these microprocessors due to faster speeds; short mean time to failures (poor reliability), etc. *Nanomaterials* help the industry break these barriers down by providing the manufacturers with nanocrystalline starting materials, ultra-high purity materials, materials with better thermal conductivity and longer-lasting and durable interconnections.

Elimination of Pollutants

Nanocrystalline materials possess extremely large grain boundaries relative to their grain size. Hence, nanomaterials are very active in terms of their chemical, physical and mechanical properties. Due to their enhanced chemical activity, nanomaterials can be used as catalysts to react with toxic gases such as carbon monoxide and nitrogen oxide in automobile catalytic converters and power generation equipment to prevent environmental pollution arising from burning gasoline and coal.

Better Insulation Materials

Nanocrystalline materials synthesised by the sol-gel technique result in foam like structures called "aerogels". These aerogels are porous and extremely lightweight; yet, they can load equivalent to 100 times their weight. Aerogels are composed of three-dimensional, continuous network of particles

with air (or any other fluid such as a gas) trapped at their interstices. Since they are porous and air is trapped at the interstices, aerogels are currently being used for insulation in offices, homes, etc. By using aerogels for insulation, heating and cooling bills are drastically reduced, thereby saving power and reducing the resultant environmental pollution. They are also being used as materials for "smart" windows, which darken when the sun is too bright (just as in changeable lenses in prescription spectacles and sunglasses) and they lighten themselves when the sun is not shining too brightly.

Low-Cost Flat-Panel Displays

Flat-panel displays represent a huge market in the computer and TV industry. By synthesising nanocrystalline phosphors, the resolution of these display devices can be greatly enhanced and the manufacturing cost can be significantly reduced. Also, the flat-panel displays constructed out of nanomaterials possess much higher brightness and contrast than the conventional ones owing to their enhanced electrical and magnetic properties.

Tougher and Harder Cutting Tools

Cutting tools made of nanocrystalline materials, such as tungsten carbide, tantalum carbide and titanium carbide are much harder, much more wear-resistant, erosion-resistant and last longer than their conventional (large-grained) counterparts. They also enable the manufacturer to machine various materials much faster, thereby increasing productivity and significantly reducing manufacturing costs. Also, for the miniaturization of microelectronic circuits, the industry needs microdrills (drill bits with diameter less than the thickness of an average human hair or 100 µm) with enhanced edge retention and far better wear resistance. Since nanocrystalline carbides are much stronger, harder and wear-resistant, they are currently being used in these microdrills.

High Energy Density Batteries

The energy density (storage capacity) of conventional batteries is quite low requiring frequent recharging. The life of conventional and rechargeable batteries is also low. Nanocrystalline materials synthesised by sol-gel techniques are candidates for separator plates in batteries because of their foam-like (aerogel) structure, which can hold considerably more energy than their conventional separators. Furthermore, nickel-metal hydride (Ni-MH) batteries made of nanocrystalline nickel and metal hydrides are envisioned to require far less frequent recharging and to last much longer because of their large grain boundary (surface) area and enhanced physical, chemical and mechanical properties.

High-Power Magnets

The strength of a magnet is measured in terms of coercivity and saturation magnetisation values. These values increase with a decrease in the grain size and an increase in the specific surface area (surface area per unit volume of the grains) of the grains. It has been shown that magnets made of nanocrystalline yttrium-samarium-cobalt grains possess very unusual magnetic properties due to their extremely large surface area. Typical applications for these high-power rare-earth magnets include quieter submarines, automobile alternators, land-based power generators, motors for ships, ultra-sensitive analytical instruments and magnetic resonance imaging (MRI) in medical diagnostics.

High-Sensitivity Sensors

The sensitivity of the sensors is based on the changes in various parameters they are designed to measure. The measured parameters include electrical resistivity, chemical activity, magnetic permeability, thermal conductivity, capacitance, etc. All these parameters depend greatly on the microstructure (grain size) of the materials employed in the sensors. A change in the sensor's environment is manifested by the sensor material's chemical, physical or mechanical characteristics, which is exploited for detection. For instance, a carbon monoxide sensor made of zirconium oxide (zirconia) uses its chemical stability to detect the presence of carbon monoxide. In the event of carbon monoxide's presence, the oxygen atoms in zirconium oxide react with the carbon in carbon monoxide to partially reduce zirconium oxide. This reaction triggers a change in the sensor's characteristics, such as conductivity (or resistivity) and capacitance. The rate and the extent of this reaction are greatly increased by a decrease in the grain size. Hence, sensors made up of nanocrystalline materials are extremely sensitive to the change in their environment. Typical applications for sensors made out of nanocrystalline materials are smoke detectors, ice detectors on aircraft wings, automobile engine performance sensor, etc.

Automobile with Better Fuel Efficiency

Currently, automobile engines waste considerable amounts of gasoline, thereby contributing to environmental pollution by not completely combusting the fuel. A conventional spark plug is not designed to burn the gasoline completely and efficiently. This problem is compounded by defective or worn-out sparks plug electrodes. Since nanomaterials are stronger, harder and much more wear-resistant and erosion-resistant, they are presently being envisioned to be used as spark plugs. These electrodes render the spark plugs longer-lasting and combust fuel far more efficiently and completely. A radically new spark plug design called the "railplug" is also in the prototype stage. This railplug uses the technology derived from the "railgun," which is a spin-off of the popular Star Wars defence program. However, these railplugs generate much more powerful sparks (with an energy density of approximately 1 kJ/mm²). Hence, conventional materials erode and corrode too soon and quite frequently to be of any practical use in automobiles. Nevertheless, railplugs made of nanomaterials last much longer than the conventional spark plugs. Also, automobiles waste significant amounts of energy by losing the thermal energy generated by the engine. This is especially true in the case of diesel engines. Hence, the engine cylinders (liners) are currently being envisioned to be coated with nanocrystalline ceramics, such as zirconia and alumina, so that they retain heat much more efficiently and result in complete and efficient combustion of the fuel.

Aerospace Components with Enhanced Performance Characteristics

Due to the risks involved in flying, aircraft manufacturers strive to make the aerospace components stronger, tougher and long lasting. One of the key properties required of the aircraft components is the fatigue strength, which decreases with the component's age. By making the components out of stronger materials, the life of the aircraft can be greatly increased. The fatigue strength increases with a reduction in the grain size of the material. Nanomaterials provide such a significant reduction in the grain size over conventional materials that the fatigue life is increased by an average of 200-300%. Furthermore, components made of nanomaterials are stronger and can operate at higher temperatures, aircraft can fly faster and more efficiently (for the same amount of aviation fuel).

In spacecraft, elevated-temperature strength of the material is crucial because the components (such as rocket engines, thrusters and vectoring nozzles) operate at much higher temperatures than aircraft and higher speeds. Nanomaterials are perfect candidates for spacecraft applications as well.

Better and Future Weapons Platforms

Conventional guns, such as cannons, 155 mm howitzers and multiple-launch rocket system (MLRS), utilize the chemical energy derived by igniting a charge of chemicals (gun powder). The maximum velocity at which the penetrator can be propelled is approximately 1.5–2.0 km/sec. On the other hand, electromagnetic launchers (EML guns), or railguns, use the electrical energy, and the concomitant magnetic field (energy), to propel the penetrators/projectiles at velocities up to 10 km/sec. This increase in velocity results in greater kinetic energy for the same penetrator mass. The greater the energy, the greater is the damage inflicted on the target. For these reasons, the U.S. Army has conducted extensive research into the railguns. Since a railgun operates on electrical energy, the rails need to be very good conductors of electricity. Also, they need to be so strong and rigid that the railgun does not sag while firing and buckle under its own weight. The obvious choice for high electrical conductivity is copper. However, the railguns made out of copper wear out much too quickly due to the erosion of the rails by the hypervelocity projectiles and they lack high-temperature strength. The wear and erosion of copper rails necessitate inordinately frequent barrel replacements. In order to satisfy these requirements, a nanocrystalline composite material made of tungsten, copper and titanium diboride is being evaluated as a potential candidate. This nanocomposite possesses the requisite electrical conductivity, adequate thermal conductivity, excellent high strength, high rigidity, hardness and wear/erosion resistance. This results in longer-lasting, wear-resistant and erosion-resistant railguns, which can be fired more frequently and often than their conventional counterparts.

Longer-Lasting Satellites

Satellites are being used for both defence and civilian applications. These satellites utilize thruster rockets to remain in or change their orbits due to a variety of factors including the influence of gravitational forces exerted by the earth. Hence, these satellites are repositioned using these thrusters. The life of these satellites, to a large extent, is determined by the amount of fuel they can carry on board. In fact, more than 1/3 of the fuel carried aboard by the satellites is wasted by these repositioning thrusters due to incomplete and inefficient combustion of the fuel, such as hydrazine. The reason for the incomplete and inefficient combustion is that the onboard ignitors wear out quickly and cease to perform effectively. Nanomaterials, such as nanocrystalline tungsten-titanium diboride-copper composite, are potential candidates for enhancing these ignitors' life and performance characteristics.

Longer-Lasting Medical Implants

Currently, medical implants, such as orthopaedic implants and heart valves, are made of titanium and stainless steel alloys. These alloys are primarily used in humans because they are biocompatible, i.e., they do not adversely react with human tissue. In the case of orthopaedic implants (artificial bones for hip, etc.), these materials are relatively non-porous. For an implant to effectively mimic a natural human bone, the surrounding tissue must penetrate the implants, thereby holding the implant

with the required strength. Since these materials are relatively impervious, human tissues do not penetrate the implants, thereby reducing their effectiveness. Furthermore these metal alloys wear out quickly necessitating frequent, and often very expensive, surgeries. However, nanocrystalline zirconia (zirconium oxide) ceramic is hard, wear-resistant, corrosion-resistant (biological fluids are corrosive) and biocompatible. Nanoceramics can also be made porous into aerogels (aerogels can withstand up to 100 times their weight), if they are synthesized by sol-gel techniques. This results in far less frequent implant replacements, and hence, a significant reduction in surgical expenses. Nanocrystalline silicon carbide (SiC) is a potential material for artificial heart valves primarily due to its low weight, high strength, extreme hardness, wear resistance inertness (SiC does not react with biological fluids) and corrosion resistance.

Ductile, Machinable Ceramics

Ceramics are usually very hard, brittle, and hard to machine. These characteristics of ceramics have discouraged the potential users from exploiting their beneficial properties. However, with a reduction in grain size, these ceramics have increasingly been used. Zirconia, a hard, brittle ceramic, has even been rendered superplastic, i.e., it can be deformed to great lengths (up to 300% of its original length). However, these ceramics must possess nanocrystalline grains to be superplastic. In fact, nanocrystalline ceramics, such as silicon nitride (Si_3N_4) and silicon carbide (SiC), have been used in such automotive applications as high-strength springs, ball bearings, and valve lifters, because they possess good formability and machinability combined with excellent physical, chemical and mechanical properties. They are also used as components in high-temperature furnaces. Nanocrystalline ceramics can be pressed and sintered into various shapes at significantly lower temperatures, whereas it would be very difficult, if not impossible, to press and sinter conventional ceramics even at high temperatures.

Large Electrochromic Display Devices

An electrochromic devices consists of materials in which an optical absorption band can be introduced or an existing band can be altered by the passage of current through the materials or by the application of an electric field. Nanocrystalline materials, such as tungstic oxide ($\text{WO}_3 \cdot x\text{H}_2\text{O}$) gel, are used in very large electrochromic display devices. The reaction governing electrochromism (a reversible coloration process under the influence of an electric field) is the double-injection of ions (or protons, H^+) and electrons, which combine with the nanocrystalline tungstic acid to form a tungsten bronze. These devices are primarily used in public billboards and ticker boards to convey information. Electrochromic devices are similar to liquid-crystal displays (LCD) commonly used in calculators and watches. However, electrochromic devices display information by changing colour when a voltage is applied. When the polarity is reversed, the colour is bleached. The resolution, brightness and contrast of these devices greatly depend on the tungstic acid gel's grain size. Hence, nanomaterials are being explored for this purpose.

Dye-Sensitized Cells

These cells employ complex dye molecules attached to the surface of nanostructured oxides like titanium (IV) oxide (TiO_2) or niobium (V) oxide (Nb_2O_5). The dyes exhibit broad light absorption profiles and rapid photo induced charge transfer of electrons to the nanocrystals. These cells show solar conversion efficiencies of about 4%.

Coatings

Nanopowders and nanoparticle dispersions have seen increasing applications in coatings. Due to their small size, very even coating can be achieved by painting nanoparticle dispersions onto a surface and baking off the residual solvent.

Optically Transparent Conductive Coatings

Indium tin oxide (ITO) and antimony tin oxide (ATO) are well-known, optically transparent, electrically conductive materials. Nanoparticles of these materials can be painted on surfaces such as interactive touch screens to create a conductive, transparent screen without relying on expensive sputtering techniques. In addition, ITO and ATO can be used as an antistatic coating, utilizing their inherent conductivity to dissipate static charge.

Resistant Coatings

Nanoscale aluminium oxide and titanium oxide are optically transparent and greatly increase the abrasion resistance of traditional coatings. Titanium oxide is of particular interest in many optical applications, since it is highly reflective for most ultraviolet radiation. Zinc oxide and rare-earth oxides are also UV-reflective, but optically transparent and are therefore effective in protecting surfaces from degradation brought about by exposure to UV radiation.

Using Biosensors and Biolabels to Understanding Living Cells

Nanotechnology has the potential to increase our ability to understand the fundamental working of living cells. Many potential applications for metal nanomaterials as biosensors and biolabels are under investigation. They have found use in cellular studies, enhanced spectroscopic techniques, biochips, and protein and enzyme analysis.

Using Fluorescent Nanoparticles for Cell Labeling and Magnetic Nanoparticles as Sensors

Multi-colour labelling of both fixed and living cells with fluorescent nanoparticles conjugated with biological ligands that specifically bind against certain cellular targets enables the recording of diffusion pathways in receptor cells. Uptake of nanoparticles into the vesicular compartments around the nucleus of cells can be used to label the cells so that their pathway and fate can be followed. The nanoparticles exhibit reduced photobleaching as compared to traditional dyes and are passed on to daughter cells during cell division, therefore allowing for much longer term observation. Magnetic nanoparticles can also act as sensors for assessing how external stressors affect changes in intracellular biochemistry and gene expression.

Diagnostics***How Nanotechnology Methods can Improve Medical Diagnostics***

Early detection of disease remains a primary goal of the medical community. Nanotechnology holds great promise for enabling the achievement of this goal. Nanoparticles in particular have exhibited tremendous potential for detecting fragments of viruses, pre-cancerous cells, disease markers and indicators of radiation damage. Gold coatings have made it possible to use toxic cobalt nanoparticles for biomedical applications. Gold coated ferromagnetic nanoparticles tagged with

HIV antibodies may be able to detect virus particles left after completion of conventional drug therapy. Metal nanoparticles in the form of dendrimers have also been functionalized with different biomolecules to detect specific proteins, antibodies and other disease indicators. Fluorescent markers can also be attached to the dendrimers.

Targeted Drug Delivery

Targeted drug delivery systems can convey drugs more effectively and/or more conveniently, increase patient compliance, extend the product life cycle, provide product differentiation and reduce health care costs. Drug delivery systems that rely on nanomaterials also allow for targeted delivery of compounds characterized by low oral bioavailability due to poor water solubility, permeability and/or instability and provide for longer sustained and controlled release profiles. These technologies can increase the potency of traditional small molecule drugs in addition to potentially providing a mechanism for treating previously incurable diseases.

Using Magnetic Nanoparticles Targeted Drug Delivery

The use of magnetic nanoparticles in targeted drug delivery systems is under investigation by several research groups. Therapeutic drug molecules have been immobilized on the surface of magnetic nanoparticles or nanocrystals and directed to a specific target tissue using a magnetic field gradient. The drug is released by applying a radio frequency (RF) pulse. Gold coated iron, nickel and cobalt ferromagnetic nanoparticles have been employed in this "tag and drag" approach. In hypothermal treatment, magnetic nanoparticles are directed to diseased tissue containing heat sensitive tumors. An a.c. magnetic field is applied such that the nanoparticles become heated, causing destruction of the cancerous cells. More effective radiation therapy for tumour treatment can also be expected using metallic nanoparticles instead of, for example, magnetite. The nanoparticles allow the application of higher dosages of radiation at the tumour while sparing normal tissue.

Other Applications for Nanomaterials in the Medical and Pharmaceutical Sector

Numerous other potential applications exist in the medical and pharmaceutical field for nanomaterials. Areas currently under investigation include gene therapy, antibacterial/antimicrobial agents for burn and wound dressings, repair of damaged retinas, artificial tissues, prosthetics, enhancing signals for magnetic resonance imaging examinations, and as radio frequency controlled switching of complex biochemical processes.

Conclusion

Nanomaterials have brought a sea change in the way materials are designed. Until the present, the quest for improved materials has had to be done indirectly, such as heat treatments for manipulating the properties of steel to give them nanoscaled structures. Nanomaterials like nanoparticles, nanotubes, nanorods, nano fibres, etc. when manufactured on large scale in industrial environment are expected to be produced in controlled manner and not pollute the environment or harm the workers. Use of nanomaterials as hydrogen storage or efficient oil filters may reduce pollution due to vehicles. Some of the nanoparticles based sensors are much more novel and sensitive compared to those being used. Nanomaterial based sensors would be smart sensors.

Besides these useful applications there are some dangers associated with the use of nanomaterials in various fields. One area of concern is the effect that industrial-scale manufacturing and use of nanomaterials would have on human health and the environment, as suggested by nanotoxicology research group such as the Center for Responsible Nanotechnology have advocated that nanotechnology should be specially regulated by governments for these reasons. Others counter that overregulation would stifle scientific research and the development of innovations which could greatly benefit mankind.

SHORT ANSWER TYPE QUESTIONS

1. What is the condition for a material to be termed nanomaterial?
2. Define the terms: (i) quantum well, (ii) quantum wire and (iii) quantum dot.
3. Why the properties of the nanomaterials are different from the bulk materials?
4. Define nanotechnology.
5. Is nanotechnology new? Discuss giving suitable examples.
6. State and explain Moore's law.
7. What are metal nanoclusters?
8. What are structural magic numbers?
9. Define density of states and explain its significance.
10. What is fullerene? Elaborate why it is also known as buckyball.
11. Briefly introduce various structures of CNT.
12. What is magnetoresistance?
13. What are Landau levels?
14. Explain the electrical properties of CNTs.
15. Briefly outline the mechanical properties of CNTs.
16. In the field of nanoelectronics, discuss various uses of carbon nanotubes.
17. What is fuel cell? How CNTs are useful for fuel cells?
18. What are supercapacitors? How CNTs are useful?
19. Explain how nanomaterials are useful for aerospace applications?
20. Explain the role of nanomaterials for medical diagnostics.

LONG ANSWER TYPE QUESTIONS

1. Discuss the impact of nanoscience and technology on the present and future technologies.
2. What are metal nanoclusters? Describe the process of their formation giving a suitable diagram.
3. Define density of states and explain its significance. Show that the density of states vary with energy and it is inversely proportional to square root of energy. Obtain expressions for density of states for OD, 1D, 2D and 3D structures.
4. Discuss different structures and properties of carbon nanotubes.
5. Discuss various synthesis methods of CNTs and mention their relative advantages and disadvantages.
6. Describe the applications of CNTs giving necessary background.
7. Mention the applications of nanomaterials in different sphere of human activity.
8. Discuss the medical applications of nanomaterials.

Index

A

A particle in a box,
wave functions of, 51 ●
Absorption coefficient,
measurement of, 230 ●
Accelerators,
AC of, 140 ●
DC of, 139 ●
Alpha decay, 110 ●
theory of, 112 ●
Amorphous solids, 211 ●
Annihilation, 27 ●
Antiferroelectricity, 318 ●
Anti-symmetric wave function, 76 ●
Artificial radioactivity, 108 ●
Asymmetry energy, 125 ●
Avalanche breakdown, 278 ●

B

Barrier potential, 271 ●
Batter fuel efficiency,
automobile with, 389 ●
BCS ground state, 351 ●
BCS theory, 348 ●
features of, 350 ●
Beta decay, 115 ●

Betatron, 144 ●

Bose-Einstein statistics, 76 ●
Bragg law,
x-ray diffraction, 213 ●
Bubble chamber, 151 ●
Bucky ball, 380 ●
Building,
acoustics of, 222 ●

C

Carbon molecules, 379 ●
Carbon nanostructures, 379 ●
Carbon nanotubes (CNTs), 381 ●
application of, 384 ●
properties of, 383 ●
synthesis of, 382 ●
Carbon-nitrogen cycle, 134 ●
Ceramics, 319 ●
Chain reaction, 128 ●
Classical mechanics,
limitations of, 3 ●
Claussius-Mossotti equation, 297 ●
Claussius-Mossotti relation, 300 ●
Cloud chamber, 151 ●
Coherence length, 353 ●
Compton effects, 21 ●

- Conducting medium,
electromagnetic waves in, 198 ●
- Conduction band, 240 ●
- Cooper-Pair formation, 349 ●
- Cooper-Pair tunneling, 355 ●
- Coulomb energy, 124 ●
- Coulomb's law, 291 ●
- Critical magnetic field, 329 ●
- Crystalline solids, 211 ●
- Current,
continuity of, 178 ●
- Cyclotron, 141 ●
- D**
- De-Broglie waves, 28 ●
- Defect breakdown, 321 ●
- Degenerate eigenfunctions, 49 ●
- Density of states, 377 ●
- Dielectric constant, 288 ●
- Dielectric dispersion, 307 ●
- Dielectric inside field, 290 ●
- Dielectric loss, 307 ●
- Dielectric materials, 286 ●
- Dielectric medium,
electromagnetic waves in, 196 ●
- Dielectric smart materials, 308 ●
- Dielectric strength, 320 ●
- Dielectrics, 291 ●
Gauss's law, 291 ●
- Diffusion currents, 260 ●
- Dihydrogen phosphate group, 315 ●
- Dimension effects, 375 ●
- Dipole moment, 288 ●
- Displacement current, 177 ●
characteristics of, 178 ●
- Drift current, 260 ●
- Dye-sensitized cells, 391 ●
- E**
- Eigenfunctions, 49 ●
- Eigenvalues, 49 ●
- Einstein's equation, 260 ●
- Einstein's photoelectric equation, 12 ●
- Electrets, 318 ●
- Electric force,
motion of charged in, 157 ●
- Electrical conductivity, 258 ●
- Electrical susceptibility, 293 ●
- Electromagnetic radiation,
absorption of, 342 ●
- Electromagnetic spectrum, 201 ●
- Electromagnetic theory, 157 ●
- Electromagnetic waves,
energy in, 188 ●
momentum in, 188 ●
- Electromechanical breakdown, 321 ●
- Electron in atom,
binding energy of, 37 ●
- Electron,
de-Broglie wavelength of, 29 ●
- Electron-hole recombination, 242 ●
- Electronic polarization, 302, 304 ●
- Electrons,
density of, 254 ●
- Electrostriction, 319 ●
- Energy band, 239 ●
- Energy bands in solids,
formation of, 239 ●
- Energy density,
growth of, 227 ●
- Energy flow, 189 ●
- Energy gap, 341 ●
- Equal probability hypothesis, 69 ●
- Equipartition of energy,
law of, 74 ●
- Expectation values, 44 ●
- Intrinsic semiconductors, 245 ●
comparison of, 248 ●
electrical conductivity of, 249 ●
types of, 245 ●

F

- Factors influencing building acoustics, 232 ●
- Fermi,
 - distribution function of, 81 ●
 - energy of, 81 ●
- Fermi-Dirac statistics, 79 ●
- Fermi-Dirac system,
 - energy of, 82 ●
- Ferroelectric hysteresis loop, 312 ●
- Ferroelectric materials, 313 ●
- Ferroelectricity, 311 ●
- Fission products,
 - mass distribution of, 130 ●
- Fission yield curve, 130 ●
- Fission,
 - neutron emission in, 131 ●
- Flux exclusion, 333 ●
- Forward biasing, 272 ●

G

- Gamma decay, 119 ●
- Gauss's law, 161 ●
 - applications of, 163 ●
 - Coulomb's law from, 162 ●
- Geiger-Muller counter, 149 ●
 - schematic diagram of, 149 ●
- Group velocity, 32 ●
- H**
- Hall coefficient,
 - determination of, 266 ●
- Hall effect, 264 ●
 - applications of, 267 ●
- High energy density batteries, 388 ●
- High-power magnets, 388 ●
- High-sensitivity sensors, 389 ●
- High-Tc superconductors, 363 ●
- High-temperature superconductors, 361 ●
- Holes,
 - density of, 256 ●

I

- Induced radioactivity, 108 ●
- Insulators, 240 ●
- Intrinsic breakdown, 320 ●
- Intrinsic carrier concentration, 257 ●
- Intrinsic carriers, 245 ●
- Intrinsic semiconductors, 245 ●
 - carrier concentration in, 254 ●
 - comparison of, 248 ●
 - Fermi level in, 258 ●
- Ionic polarization, 303 ●
- Ionization breakdown, 321 ●
- Ionization chamber, 147 ●
- Isotope effect, 343 ●

J

- Jellium model, 375 ●
- Josephson effects, 355 ●
- Junction diode,
 - resistance of, 275 ●

K

- Kammerling ones experiment, 325 ●

L

- Large electrochromic display devices, 391 ●
- Linear accelerator, 140 ●
- Liouville's theorem, 68 ●
- Liquid drop model, 123 ●
- Local electric field, 297 ●
- London equation, 344 ●
- Longer-lasting satellites, 390 ●
- Lorentz force, 159 ●
- Lorentz relation, 300 ●
- Loss-less power transmission, 366 ●
- Low-cost flat-panel displays, 388 ●

M

- Macroscopic quantum interference, 359 ●

- Macroscopic systems, 67 ●
Magnetic field,
 effect of, 329 ●
Magnetic flux quantization, 353 ●
Magnetic focussing, 160 ●
Magnetic force,
 motion of charged in, 157 ●
Magnetic lens, 160 ●
Magnetic resonance imaging, 336 ●
Magnetostriction ultrasonic generator, 218 ●
Matter waves, 28 ●
 verification of, 30 ●
Maxwell's displacement current, 176 ●
Maxwell's equations, 182 ●
Maxwell's speed distribution law, 71 ●
Maxwell-Boltzmann distribution law, 70 ●
Meissner effect, 333 ●
Metal nanoclusters, 373 ●
Microscopic systems, 67 ●
Minority carriers,
 generation of, 263 ●
 recombination of, 263 ●
Molecular orbital theory, 375 ●
Momentum flow, 191 ●
Moore's laws, 372 ●
- N**
- Nano science, 371 ●
Nanomaterials, 370 ●
 application of, 387 ●
Nanoparticles, 372 ●
Next-generation computer chips, 387 ●
Non-degenerate eigenfunctions, 49 ●
Non-ohmic device, 275 ●
N-type semiconductor, 246 ●
 comparison of, 248 ●
Nuclear binding energy, 96 ●
Nuclear fission, 128 ●
Nuclear forces, 93 ●
 Meson theory of, 94 ●
- Nuclear fusion, 133 ●
Nuclear holocaust, 135 ●
Nuclear models, 123 ●
Nuclear physics, 89 ●
Nuclear radiation detectors, 147 ●
Nuclear reactions, 120 ●
Nuclear reactor, 131 ●
Nuclei,
 stability of, 96 ●
Nucleus,
 general properties of, 90 ●
 structure of, 89 ●
- O**
- One-dimensional structure, 378 ●
Optimum reverberation time, 230 ●
Orientational polarization, 302 ●
- P**
- Pair production, 9, 27 ●
Pairing energy, 127 ●
Penetration depth, 337 ●
Perovskites group, 315 ●
Persistent current, 328 ●
Phase space, 68 ●
Phase velocity, 32 ●
Photoelectric cells, 14 ●
 applications of, 15 ●
Photoelectric effect, 9 ●
Photoelectric emission,
 explanation of, 12 ●
Photoemissive cell, 14 ●
Photons, 7 ●
Phototube, 9 ●
Piezoelectric ultrasonic generator, 217 ●
Piezoelectricity, 308 ●
Piezoelectrics,
 classification of, 310 ●
Plank's radiation law, 3 ●

- P-N junction diode,
barrier potential in, 271 ●
biasing of, 272 ●
characteristics of, 273 ●
depletion layer in, 271 ●
formation of, 270 ●
Polarizability, 293 ●
Polarization, 288 ●
types of, 302 ●
Poynting vector, 189 ●
Poynting's theorem, 192 ●
Proportional counter operation, 149 ●
Proton-proton cycle, 134 ●
P-type semiconductor, 247 ●
comparison of, 248 ●
Pyroelectricity, 311 ●
- Q**
Quantum dot, 370 ●
Quantum mechanics,
bra-ket notation in, 43 ●
operators in, 45 ●
Quantum statistics, 76 ●
Quantum structures, 375 ●
Quantum theory, 3 ●
Quantum well, 370, 376 ●
Quantum wire, 370 ●
- R**
Radiation hazards, 135 ●
Radiation pressure, 191 ●
Radiation,
dual nature of, 28 ●
Radioactive decay law, 101 ●
Radioactivity,
applications of, 109 ●
Rayleigh-Jean's law, 7 ●
Recoil electron,
direction of, 23 ●
kinetic energy of, 24 ●
- Rectangular potential barrier, 55 ●
Reverberation time,
importance of, 230 ●
Reverberation, 223 ●
Reverse biasing, 272 ●
Reverse current, 273 ●
Rochelle salt, 314 ●
- S**
Sabine's experiments, 223 ●
Sabine's formula,
derivation of, 224 ●
Schrödinger equation, 47, 48 ●
Scintillation detectors, 150 ●
Semiconductors, 240 ●
characteristics of, 249 ●
current carriers in, 243 ●
Shell model, 127 ●
Simple harmonic oscillator, 52 ●
Single particle,
tunneling of, 354 ●
Sinusoidal electromagnetic waves, 186 ●
Solid state detectors, 150 ●
Space-charge polarization, 303 ●
Spaced energy level, 239 ●
Statistical equilibrium, 69 ●
Statistical mechanics, 67 ●
Steady state form, 48 ●
Superconducting fullerides, 364 ●
Superconducting magnets, 366 ●
Superconductivity, 325 ●
discovery of, 325 ●
Superconductors,
A.C. resistivity in, 328 ●
applications of, 365 ●
magnetic phase diagram of, 330 ●
types of, 335 ●
Surface energy, 124 ●
Symmetric wave function, 76 ●

T

- Target drug delivery, 393 ◉
- Thermal breakdown, 321 ◉
- Thermal conductivity, 340 ◉
- Thermal neutrons, 129 ◉
- Thermal properties, 338 ◉
- Thermal switches, 340 ◉
- Three-dimensional structure, 378 ◉
- Time dependent form, 47 ◉
- Tin,
 - superconducting transition in, 327 ◉
- Transmission probability, 58 ◉
- Tunnel diode parameters, 282 ◉
- Tunnel diode, 280 ◉
 - applications of, 283 ◉
 - advantages of, 283 ◉
 - disadvantages of, 283 ◉
 - V-I characteristic of, 281 ◉
- Tunnel effect, 55 ◉
- Tunneling process, 282 ◉
- Two-dimensional structure, 378 ◉

U

- Ultrasonic waves, 217 ◉
 - applications of, 220 ◉
 - detection of, 220 ◉
 - production of, 217 ◉
- Uncertainty principle, 35 ◉
- Unstable nuclei,
 - decay of, 100 ◉

Uranium,

- isotopes of, 130 ◉

V

- Vacuum,
 - plane electromagnetic wave in, 183 ◉
- Valence band, 240 ◉
- Voltage regulator, 277 ◉
- Volume energy, 123 ◉

W

- Wave equation, 183 ◉
- Wave function, 42 ◉
- Wave packet, 32 ◉
- Weizsäcker semiempirical formula, 123 ◉
- Wien's displacement law, 6 ◉

X

- X-ray spectra, 208 ◉
- X-rays,
 - introduction of, 207 ◉
 - production of, 207 ◉
 - properties of, 207 ◉

Z

- Zener breakdown, 278 ◉
- Zener diode, 277 ◉
- Zero dimensional structure, 377 ◉
- Zero resistance, 328 ◉