Physics for Electronics Engineering and Information Science

About the Authors



S Salivahanan is the Principal of SSN College of Engineering, Chennai. He obtained his B.E. degree in electronics and communication engineering from PSG College of Technology, Coimbatore, M.E. degree in communication systems from NIT, Trichy, and Ph.D. in the area of microwave integrated circuits from Madurai Kamaraj University. He has four decades of teaching, research, administration and industrial experience, both in India and abroad. He has taught at NIT, Trichy, A.C. College of Engineering and

Technology, Karaikudi, R.V. College of Engineering, Bangalore, and Mepco Schlenk Engineering College, Sivakasi. He has industrial experience as scientist/engineer at Space Applications Centre, ISRO, Ahmedabad, Telecommunication Engineer at State Organization of Electricity, Iraq, and Electronics Engineer at Electric Dar Establishment, Kingdom of Saudi Arabia.

He is the author of 40 popular books which include all-time bestsellers such as *Basic Electrical and Electronics Engineering, Electronic Devices and Circuits, Linear Integrated Circuits* published by McGraw Hill Education (India), *Digital Signal Processing* by McGraw Hill International, which has also been translated into Mandarin, the Chinese language. He has also authored the books on *Digital Circuits and Design, Electromagnetic Field Theory, Circuit Theory, Network Analysis and Synthesis* and *Control Systems Engineering*. He has published several papers at national and international levels.

Professor Salivahanan is the recipient of Bharatiya Vidya Bhavan National Award for Best Engineering College Principal for 2011 from ISTE, and IEEE Outstanding Branch Counsellor and Advisor Award in the Asia-Pacific region for 1996–97. He was the Chairman of IEEE Madras Section for 2008–2009 and Syndicate Member of Anna University.

He is a Senior Member of IEEE, Fellow of IETE, Fellow of Institution of Engineers (India), Life Member of ISTE and Life Member of Society for EMC Engineers. He is also a member of IEEE societies in Microwave Theory and Techniques, Communications, Signal Processing, and Aerospace and Electronics.



Dr. A Rajalakshmi is a Professor and Head at the Department of Physics, SSN College of Engineering, Chennai. She obtained her M.Sc. degree in nuclear physics from University of Madras. She won the K. S. Krishnan Gold Medal, The Professor Subramanya Ayyar Commemoration Medal, The Jagirdar of Arni Medal, Dr. K. S. Krishnan Memorial Price and The Smt. K. Vimala Devi Medal, for securing first rank in her M.Sc. degree programme. She obtained her Ph.D. in Physics from IIT Madras, in the field of

X-ray spectroscopy. She has more than 30 years of experience in teaching and research. Her current field of research is environmental and radiation physics. She has published several papers in peer reviewed journals and conferences. She has been a member of the Physics Syllabus Sub-committee for the B.E./B.Tech. programmes of Anna University several times. She is also a Life Member of ISTE.



S Karthie is working as an Assistant Professor at the Department of ECE, SSN College of Engineering, Chennai, and has over 14 years of teaching experience. Currently, he is pursuing research in the area of microwave filter design. He has published several research papers in National and International Conferences. He has also co-authored a book on *Electromagnetic Field Theory*. His areas of interest include electromagnetics, electronic devices and circuits, RF and microwave integrated circuits. He is an active

member of IEEE – Microwave Theory and Techniques Society. He is also a Life member of IETE and ISTE.



Dr. N P Rajesh is working as an Assistant Professor at the Department of Physics and Head of Centre for Radiation, Environmental Science and Technology, SSN College of Engineering, Chennai. He obtained his doctorate from Anna University, Chennai, in the field of crystal growth and characterization of nonlinear optical crystals. He has two decades of research and teaching experience, and published 90 papers in refereed international journals. He has received Young Scientist Award from International organisation of

Crystal Growth, Germany and DST, Government of India, AICTE Career Award for Young Teachers, and DAIKO Foundation Research Fellowship, Japan. He is the fellow of Optical Society of India (OSI), Life Member of Indian Science Congress Association (ISCA) and Indian Society for Technical Education (ISTE), and Member of Indian Society for NDT (ISNDT).

Physics for Electronics Engineering and Information Science

S SALIVAHANAN Principal

A RAJALAKSHMI

Professor and Head, Department of Physics S KARTHIE

Assistant Professor, Department of ECE

N P RAJESH Assistant Professor, Department of Physics

SSN College of Engineering, Chennai, Tamil Nadu



McGraw Hill Education (India) Private Limited CHENNAI

McGraw Hill Education Offices

Chennai New York St Louis San Francisco Auckland Bogotá Caracas Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal San Juan Santiago Singapore Sydney Tokyo Toronto



ucation McGraw Hill Education (India) Private Limited

Published by McGraw Hill Education (India) Private Limited

444/1, Sri Ekambara Naicker Industrial Estate, Alapakkam, Porur, Chennai 600 116

Physics for Electronics Engineering and Information Science

Copyright © 2018 by McGraw Hill Education (India) Private Limited.

No part of this publication may be reproduced or distributed in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise or stored in a database or retrieval system without the prior written permission of the publishers. The program listings (if any) may be entered, stored and executed in a computer system, but they may not be reproduced for publication.

This edition can be exported from India only by the publishers, *McGraw Hill Education (India) Private Limited*.

1 2 3 4 5 6 7 8 9 D102739 22 21 20 19 18

Printed and bound in India.

ISBN (13): 978-93-87572-46-1 ISBN (10): 93-87572-46-3

Vice President—Finance & Operations: Ashutosh Verma

Director—Science & Engineering Portfolio: *Vibha Mahajan* Senior Portfolio Manager—Science & Engineering: *Hemant K Jha* Associate Portfolio Manager—Science & Engineering: *Vaishali Thapliyal*

Production Head: Satinder S Baveja Copy Editor: Taranpreet Kaur Assistant Manager—Production: Anuj K Shriwastava

General Manager—Production: Rajender P Ghansela Manager—Production: Reji Kumar

Information contained in this work has been obtained by McGraw Hill Education (India), from sources believed to be reliable. However, neither McGraw Hill Education (India) nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw Hill Education (India) nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw Hill Education (India) and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Typeset at Text-o-Graphics, B-1/56, Aravali Apartment, Sector-34, Noida 201 301, and printed at

Cover Printer:

Visit us at: www.mheducation.co.in

Preface

Physics for Electronics Engineering and Information Science is designed specifically to cater to the needs of second semester B.E./B.Tech. in ECE, EEE, CSE, IT, Medical Electronics, BME, Computer and Communication Engineering, E&I and ICE students. The book has a perfect blend of focused content and complete syllabus coverage. Solved university question papers, tagged with specific topics, will be extremely helpful to students from the examination point of view. Simple, easy-to-understand and difficult-jargon-free text elucidates the fundamentals of Physics. Several solved examples, schematic diagrams and adequate questions further help students to understand and apply the concepts.

Salient Features

- Crisp content strictly as per the latest syllabus of Physics for Electronics Engineering and Physics for Information Science
- Comprehensive coverage with lucid content presentation
- Solutions to examination papers present appropriately within the chapters
- Rich exam-oriented pedagogy
 - Solved Numerical Examples within chapters
 - > Two Mark Questions and Answers at the end of each chapter
 - Unsolved Review Questions

Chapter Organisation

Chapter 1 deals with electrical properties of materials, classical free electron theory, electrical conductivity, thermal conductivity, Wiedemann-Franz law, Fermi-Dirac statistics, density of energy states, electron in periodic potential, Bloch theorem, energy bands in solids, tight binding approximation, electron effective mass and concept of hole.

Chapter 2 is devoted to semiconductor physics, intrinsic semiconductor, energy band diagram, direct and indirect semiconductor, extrinsic semiconductor, variation of carrier concentration with temperature, variation of Fermi level with temperature and impurity concentration, carrier transport, velocity-electric field relations, drift and diffusion transport, Einstein's relation, Hall effect devices, Zener and avalanche breakdown in *PN* junctions, Ohmic contacts, tunnel diode, Schottky diode, MOS capacitor and power transistor.

vi Preface

Chapter 3 describes the magnetic and dielectric properties of materials, magnetism in materials, magnetization, magnetic permeability and susceptibility, types of magnetic materials, microscopic classification of magnetic materials, ferromagnetism: origin and exchange interaction, saturation magnetization and Curie temperature, Domain theory, hard and soft magnetic materials, magnetic principle in computer data storage, magnetic hard disk (GMR sensor), dielectric materials, polarization processes, dielectric loss, internal field, Clausius-Mosotti relation, electric breakdown, high-*k* dielectrics.

Chapter 4 focuses on optical properties of materials, classification of optical materials, carrier and recombination processes, absorption, emission and scattering of light in metals, insulators and semiconductors, photocurrent in *PN* diode, solar cell, photo detectors, LED, organic LED, laser diodes, excitons, quantum confined Stark effect, quantum dot laser and optical data storage techniques.

Chapter 5 concentrates on nanoelectronic devices, electron density in bulk material, size dependence of Fermi energy, quantum confinement, quantum structures, density of states in quantum well, quantum wire and quantum dot structures, Zener-Bloch oscillations, resonant tunneling, quantum interference effects, mesoscopic structures, conductance fluctuations and coherent transport, Coulomb blockade effects, single electron phenomena and single electric transistor, magnetic semiconductors, spintronics, band gap of nanomaterials, tunneling, conductivity of metallic nanowires, ballistic transport, quantum resistance and conductance, and properties and applications of carbon nanotubes.

Acknowledgements

We sincerely thank the management of SSN College of Engineering, Chennai, for their constant encouragement, and for providing necessary facilities for completing this project. We are thankful to our colleagues, Mr. R. Gopalakrishnan, Mr. A. Chakrapani, Mr. K. Rajan and Mr. M. Suresh, for word processing the manuscript.

We are highly appreciative of the editorial and production teams at McGraw Hill Education (India), for their initiation and support to bring out this edition in a short span of time.

We would also like to take the opportunity to thank the reviewers, especially our colleagues, Dr. Prita Nair, Dr. A. Chandrasekaran, Dr. Julie Charles and Mr. S. Singaravadivelu from the Department of Physics, Dr. R. Kalidoss, Dr. K. Vishvaksenan and Dr. K.K. Nagarajan from the Department of Electronics and Communication Engineering, Dr. T. Sree Sharmila and Dr. S. Sasi Rekha from the Department of Information Technology, for their useful comments and suggestions.

Finally, we thank our family members, including our spouses, children and parents, for their patience and constant inspiration during the preparation of this book.

Any constructive criticism, suggestions and corrections for further improvement of the book will be most appreciated.

S Salivahanan A Rajalakshmi S Karthie N P Rajesh

Contents

Preface					
1.	. Electrical Properties of Materials 1.3				
	1.1	Introduction 1.1			
	1.2	Electron Theory of Solids 1.1			
	1.3	Postulates of Classical Free Electron Theory 1.3			
	1.4	Electrical Conductivity 1.4			
	1.5	Thermal Conductivity 1.13			
	1.6	Wiedemann-Franz Law and Lorentz Number 1.15			
	1.7	Success and Failures of Classical Free Electron Theory 1.17			
	1.8	Quantum Free Electron Theory 1.19			
1.9 Electrons in a Metal-Particle in a Three-Dimensional Box 1.19					
	1.10	Density of Energy States 1.27			
	1.11	Fermi-Dirac Statistics 1.31			
	1.12	Work Function 1.40			
	1.13	Band Theory of Solids 1.44			
	1.14	Electron in a Periodic Potential 1.44			
	1.15	Classification of Solids–Metals, Semiconductors and Insulators 1.49			
	1.16	Energy Bands in Solids – Tight Binding Approximation 1.50			
	1.17	Effective Mass of Electron 1.53			
	1.18	Concept of Hole 1.55			
	Two Mark Questions and Answers 1.56				
	Review Questions 1.61				
2. Semiconductor Physics		2.1-2.102			
	2.1	Introduction 2.1			
	2.2	Atomic Energy Level Diagram 2.1			
	2.3	Practical Semiconductor Materials 2.2			
	2.4	Classification of Semiconductors 2.3			
	2.5	Intrinsic Semiconductors with Energy Band Diagram 2.4			
	2.6	Direct and Indirect Bandgap Semiconductors 2.5			
	2.7	Conductivity of Semiconductor 2.6			
	2.8	Carrier Concentration in Intrinsic Semiconductors 2.7			

viii Contents

- 2.9 Properties of Intrinsic Semiconductors 2.20
- 2.10 Extrinsic Semiconductors with Energy Band Diagram 2.20
- 2.11 Carrier Concentration in N-Type and P-Type Semiconductors 2.23
- 2.12 Variation of Carrier Concentration with Temperature 2.28
- 2.13 Variation of Fermi Level with Temperature and Impurity Concentration 2.31

3.1-3.81

- 2.14 Variation in Semiconductor Parameters with Temperature 2.32
- 2.15 Carrier Transport: Velocity-Electric Field Relations 2.43
- 2.16 Drift and Diffusion Transport 2.45
- 2.17 Einstein Relationship for Semiconductor 2.48
- 2.18 Forward and Reverse Bias Characteristics of PN Junction Diode 2.49
- 2.19 Energy-Band Structure of Open-Circuited PN Junction 2.56
- 2.20 Diode Current Equation 2.58
- 2.21 Zener and Avalanche Breakdown in PN Junctions 2.59
- 2.22 Hall Effect and Devices 2.60
- 2.23 Metal-Semiconductor Junctions 2.67
- 2.24 Ohmic Contacts 2.71
- 2.25 Schottky Diode 2.73
- 2.26 Tunnel Diode 2.74
- 2.27 MOS Capacitor 2.77
- 2.28 Power Transistor 2.81
- Two Mark Questions and Answers 2.87
- Review Questions 2.98

3. Magnetic and Dielectric Properties of Materials

- 3.1 Introduction 3.1
- 3.2 Magnetism in Materials 3.1
- 3.3 Magnetic Field and Induction 3.4
- 3.4 Magnetization 3.4
- 3.5 Magnetic Permeability and Magnetic Susceptibility 3.5
- 3.6 Types of Magnetic Materials 3.9
- 3.7 Ferromagnetism Origin and Exchange Interaction 3.20
- 3.8 Saturation Magnetization and Curie Temperature 3.22
- 3.9 Domain Theory of Ferromagnetism 3.23
- 3.10 Hard and Soft Magnetic Materials Examples and Uses 3.28
- 3.11 Magnetic Principle in Computer Data Storage 3.32
- 3.12 Magnetic Hard Disk (GMR Sensor) 3.32
- 3.13 Dielectrics 3.36
- 3.14 Dielectric Constant 3.37
- 3.15 Polarization An Atomic View 3.38
- 3.16 Dielectric Polarization 3.38
- 3.17 Polarization Processes Types of Polarization 3.44
- 3.18 Local or Internal Field 3.52
- 3.19 Clausius-Mosotti Relation 3.55

Contents ix

3.20 Frequency Dependence of Polarization 3.58 3.21 Effect of Temperature on Dielectric Constant 3.59 3.22 Dielectric Loss 3.60 3.23 Dielectric Breakdown 3.62 3.24 High-k Dielectrics-Ferroelectricity 3.64 Two Marks Questions and Answers 3.68 Review Questions 3.79 4. Optical Properties of Materials Introduction 4.1 4.1 4.2 Light and Electromagnetic Radiation 4.1 4.3 Spectral Response of Human Eye 4.4 4.4 Light Interaction with Materials 4.5 4.5 Classification of Optical Materials 4.5 4.6 Carrier Generation and Recombination Process 4.7 4.7 Absorption, Emission and Scattering of Light in Metals, Insulators and Semiconductors 4.8 4.8 Photodetectors 4.13 4.9 Photoconductor 4.14 4.10 Photocurrent in a *PN* Photodiode 4.17 4.11 PIN Photodiode 4.18 4.12 Avalanche Photodiode (APD) 4.19 4.13 Phototransistor 4.19 4.14 Photodetector as Optocoupler 4.21 4.15 Photovoltaic Cell 4.22 4.16 Solar Cell *4.24* 4.17 Light Emitting Diode (LED) 4.24 4.18 Organic LED (OLED) 4.26

- 4.19 Laser Diode 4.29
- 4.20 Excitons 4.30
- 4.21 Quantum Confined Stark Effects (QCSE) 4.32
- 4.22 Quantum Dot Lasers 4.32
- 4.23 Optical Data Storage Techniques 4.33
- *Two Mark Questions and Answers* 4.38

Review Questions 4.43

5. Nanoelectronic Devices

- 5.1 Introduction 5.1
- 5.2 Electron Density in Bulk Materials 5.2
- 5.3 Size Dependence of Fermi Energy 5.2
- 5.4 Bandgap of Nanomaterials 5.3
- 5.5 Quantum Confinement 5.4
- 5.6 Quantum Structures 5.6
- 5.7 Density of States in Quantum Structures 5.7

5.1-5.42

4.1-4.44

- x Contents
 - 5.8 Zener-Bloch Oscillations 5.13
 - 5.9 Resonant Tunneling 5.14
 - 5.10 Quantum Interference Effects 5.16
 - 5.11 Mesoscopic Structures: Conductance Fluctuation and Coherent Transport 5.17
 - 5.12 Conductivity of Metallic Nanowires 5.17
 - 5.13 Ballistic Transport 5.19
 - 5.14 Quantum Resistance and Conductance 5.21
 - 5.15 Coulomb Blockade Effects 5.22
 - 5.16 Single Electron Phenomena and Single Electron Transistor 5.23
 - 5.17 Magnetic Semiconductors 5.27
 - 5.18 Spintronics 5.27
 - 5.19 Carbon Nanotubes 5.32
 - 5.20 Properties of CNT 5.34
 - 5.21 Applications of CNT 5.35
 - Two Mark Questions and Answers 5.37

Review Questions 5.41



Electrical Properties of Materials

1.1 INTRODUCTION

The field of materials science plays an important role in the understanding of various properties of materials like physical, chemical, electrical, mechanical, magnetic etc. All these have a wide range of applications in the field of engineering, medicine and day to day life. Of these, many materials, by virtue of their electrical properties, find applications in electrical appliances, solar cells, capacitors, transformers, microwave oven etc.

These are the results of various theories developed by scientists to correlate with the experimentally observed values. This, in fact, has led to the improvement of material quality and also invention of new materials. Hence, the electrical properties of materials with theories are explained in a chronological manner in this chapter.

The characteristic of conducting materials is high electrical conductivity due to the presence of free electrons. Hence, good conducting materials are used for the conduct of electricity.

Most metals and some non-metals are electrical conductors and their conducting properties vary widely depending on the free electrons. In this context, conductance can be defined as the property of the material by which it allows flow of electric current. Conductance is the reciprocal of resistance and its unit is mho (\mho).

Based on electrical conductivity, materials can be classified as (i) Good conductors (ii) Semiconductors and (iii) Nonconductors or Insulators

The electron theory of solids was developed to explain the properties of the above mentioned materials.

1.2 ELECTRON THEORY OF SOLIDS

The scientists have developed numerous theories like classical free electron theory, quantum free electron theory, band theory etc., to correlate the theory and experimental values to explain the structure and properties of solids as follows:

Classical Free Electron Theory

Drude and Lorentz developed the classical free electron theory in 1900. This assumes that the metals contain free electrons which behave like the molecules of a perfect gas. They obey the Maxwell-Boltzmann statistics and laws of classical mechanics. They are also assumed to move in a constant potential.

Based on classical free electron theory, Ohm's law, Wiedemann-Franz law (i.e ratio of electrical conductivity to the thermal conductivity) etc. can be explained satisfactorily. It fails to explain many experimentally observed facts. For example, the variation of electrical conductivity of a metal with temperature cannot be explained. Experimentally, it is known that specific heat of all solids including metal is same whereas on the basis of free electron theory, there should be a difference between the specific heat of metal and non-metal at room temperature.

Quantum Free Electron Theory

To remove the shortcomings of the free electron theory, it was modified by Pauli and Sommerfeld in 1928. They suggested that Fermi Dirac statistical law should be applied to the electron gas rather than the Maxwell-Boltzmann statistical law. The free electrons obey quantum laws and the wave nature of electrons is taken into account. The free electrons are assumed to move in a constant potential and the Fermi level electrons are responsible for determining the properties of materials. With this modification, the theory can explain all the phenomena which were not explained earlier. But this theory is still not able to explain semiconductors and insulators, the positive value of Hall coefficient etc.

Band Theory

This was developed by Bloch in 1928 and in 1931 Kronig and Penny used a simplified model to solve Schrödinger's equation for particle in a box. Here, the electrons are assumed to move in a periodic potential. The Quantum free electron theory was modified by taking into consideration the interaction of free electrons with the residual positive ions in the metal. Due to this interaction, the electrons in metals as well as in nonmetals can occupy only discrete set of energy levels which are grouped into separate energy bands. The electrons in metals can possess only those energies which lie in one of the allowed energy bands. Introducing this idea, the variation of electrical conductivity with temperature, superconductivity and many other experimental observations like classification of materials into conductors, semiconductors and insulators can be explained. The origin of band gap, concept of hole and the effective mass of electrons are the special features of this theory.

Before proceeding to classical free electron theory, some basic definitions and microscopic form of Ohm's law are to be explained.

Microscopic form of Ohm's Law

The electrical conductivity of the material is defined as the ratio of current density 'J' in (A/m^2) to the applied field intensity 'E' in (V/m).

i.e..

$$\sigma = \frac{J}{E} \tag{1.1}$$

where σ is the electrical conductivity in mhos/m or Siemens/m.

i.e..

$$J = \sigma E \tag{1.2}$$

The above equation is the *microscopic form of Ohm's law*.

The conductivity σ depends on the nature of the material and not on its dimensions. It also depends on temperature and impurity concentrations in the material. The electrical conductivity of the material is given by

$$\sigma = \frac{1}{\rho} \tag{1.3}$$

where ρ is the resistivity in $\Omega \cdot m$.

Also,

$$\sigma = \frac{l}{RA} \tag{1.4}$$

where *l* is the length of the conductor in metre (m), *R* is the resistance of the conductor in Ohm (Ω) and *A* is the area of cross section of the conductor in m². The unit of σ is $\frac{1}{\Omega \cdot m}$ or $(\Omega \cdot m)^{-1}$. It is also known that

$$J = \frac{I}{A} \tag{1.5}$$

$$E = \frac{V}{l} \tag{1.6}$$

and

where I is the current passing through the conductor in ampere and V is the voltage applied across the length of the conductor in volts.

Hence, from Eqns. (1.4), (1.5) and (1.6), we have

$$\sigma = \frac{l}{RA} = \frac{J}{E} = \frac{I/A}{V/l}$$
$$\frac{l}{RA} = \frac{I/A}{V/l}$$
$$V = RI$$
(1.7)

i.e.,

The above equation is the well known macroscopic form of Ohm's law.

1.3 POSTULATES OF CLASSICAL FREE ELECTRON THEORY

- (i) In a metal, there are fixed ion cores with the valence electrons of atoms free to move about the whole volume like the molecules of a perfect gas in a container. This can be visualized as an array of atoms (ions) permeated by a gas of free electrons as shown in Fig. 1.1. The mutual repulsion between the electrons is neglected.
- (ii) In the absence of field, the electrons move in random directions, making collisions from time to time with the positive ion cores and other free electrons as shown in Fig. 1.2. As the motion is random, the resultant velocity in any particular direction is zero.
- (iii) The velocities of electrons obey classical Maxwell Boltzmann distribution of velocities.





Fig. 1.2 Random motion of electrons

1.4 Physics for Electronics Engineering and Information Science

- (iv) The free electron gas is considered to obey kinetic theory of gases.
- (v) When a potential difference *V* is applied across the metals, the free electrons slowly drift opposite to the field, but continue to collide with the positive ion cores fixed in the lattice.
- (vi) They finally acquire a velocity called *drift velocity* v_d which is the average velocity acquired by the electron opposite to the direction of the electric field *E* as shown in Fig. 1.3.
- (vii) A relaxation time τ can be defined here as the time taken by the free electron to reach its equilibrium position from its disturbed position in the presence of the electric field.
- (viii) The collision time τ_c is the average time taken by a free electron between two successive collisions. In case of metals, $\tau = \tau_c$.



1.4 ELECTRICAL CONDUCTIVITY

To derive the expression for electrical conductivity, the drift velocity of the electron has to be determined first. In order to calculate the drift velocity, consider a free electron in an electric field 'E'.

The force experienced by an electron in the electric field is

a

$$F = eE \tag{1.8}$$

where *e* is the charge of an electron.

This accelerates the electron and according to the Newton's second law, the force is

$$F = ma \tag{1.9}$$

where *a* is the acceleration of the electron in m/s^2 and *m* is the mass of the electron in kg. From Eqns. (1.8) and (1.9), we have

$$ma = eE \tag{1.10}$$

i.e.,

$$=\frac{eE}{m} \tag{1.11}$$

It appears that the electrons should be accelerated continuously as a result of the electric field. But this is not the case.

During their motion, the electrons collide with phonons (one quantum of lattice vibration is called a *phonon*), impurities, lattice imperfections and finally with the positive ion core. Due to this, they regularly lose their kinetic energy and hence, the velocity they gained in the field. The electrons, in other words, have to surmount an opposing force F_r during their motion through the lattice. This force is proportional to the drift velocity and is directed against it as given by

$$F_r = -\frac{mv'_d(t)}{\tau} \tag{1.12}$$

where τ is the relaxation time in seconds.



(1.15)

(1.22)

Hence, the equation for directional motion of the electron may be written as

 $v'_d(t) = a \operatorname{constant} = v_d$

$$m\frac{dv'_d(t)}{dt} = eE - \frac{mv'_d(t)}{\tau}$$
(1.13)

The velocity of directional motion of electrons will rise and the electron will be accelerated until the two forces on the right-hand side of Eqn. (1.13) become equal. This is when the drift velocity has attained its maximum value. The resultant force acting on the electron and hence, its acceleration become zero.

This implies that,

$$m\frac{dv'_d(t)}{dt} = 0 \quad \text{or} \quad \frac{dv'_d(t)}{dt} = 0 \tag{1.14}$$

i.e.,

i.e.,

where v_d is the average drift velocity.

This implies that

$$eE = \frac{mv_d}{\tau} \tag{1.16}$$

$$v_d = \frac{eE\tau}{m} \tag{1.17}$$

Now, the current density (i.e., the number of charges flowing per unit area per sec) can be written as

$$J = (ne)v_d \tag{1.18}$$

where $n = \frac{\text{Number of free electrons}}{\text{m}^3}$.

Substituting Eqn. (1.17) in Eqn. (1.18), we get

$$J = \left(\frac{neeE\tau}{m}\right)$$
$$J = \left(\frac{ne^{2}\tau}{m}\right)E$$
(1.19)

i.e.,

The above equation is exactly similar to the electrical conductivity of metals in Eqn. (1.2) i.e. $J = \sigma E$. Comparing Eqn. (1.2) and Eqn.(1.19), we get

$$\sigma = \frac{J}{E} = \left(\frac{ne^2\tau}{m}\right) \tag{1.20}$$

A quantity called mobility of the electron is defined as the drift velocity per unit field.

 $v_d = \mu E$

i.e.,
$$\mu = \left(\frac{v_d}{E}\right) \tag{1.21}$$

Since $J = nev_d$, we may write

$$I = (ne\mu)E \tag{1.23}$$

$$\sigma = ne\mu$$
 (since $J = \sigma E$) (1.24)

Comparing Eqns. (1.20) and (1.24), we have

$$\mu = \frac{e\tau}{m} \tag{1.25}$$

Hence, the equations for conductivity σ and resistivity ρ are summarized as

$$\sigma = ne\mu = \left(\frac{ne^2\tau}{m}\right) \tag{1.26}$$

and

$$\rho = \frac{1}{\sigma} = \frac{1}{ne\mu} = \left(\frac{m}{ne^2\tau}\right) \tag{1.27}$$

It is assumed that the field is cut off as soon as the velocity of directional motion (drift) attains a constant value v_d . This velocity of drift starts reducing due to collisions of electrons with phonons, impurities, lattice imperfections, ion cores etc., and ultimately returns to its equilibrium state (random motion). This is known as *relaxation process*.

When the field is cutoff, E = 0 and Eqn.(1.13) becomes

$$\frac{dv'_d(t)}{dt} = -\frac{v'_d(t)}{\tau}$$

$$\frac{dv'_d(t)}{v'_d(t)} = -\frac{dt}{\tau}$$

$$\frac{dv'_d(t)}{v'_d(t)} = -\frac{dt}{\tau}$$
(1.28)

Integrating, we obtain

$$\ln v'_d(t) = -\frac{t}{\tau} + \ln v'_d(0)$$
(1.29)

(1.31)

Taking exponential on both the sides, we have

$$e^{\ln v'_d(t)} = e^{\left[-\frac{t}{\tau} + \ln v'_d(0)\right]}$$

Therefore,

$$v'_d(t) = v'_d(0) e^{(-t/\tau)}$$
 (since $e^{(a+b)} = e^{(a)} \cdot e^{(b)}$) (1.30)

Here, $v'_d(0) = v_d$ is the steady state drift velocity of an electron at the instant *E* is cut-off. Since τ is very small i.e., 10^{-14} s, $v'_d(t)$ relaxes to zero very rapidly. A graph of $v'_d(t)$ versus *t* is shown in Fig. 1.4.

Substituting $t = \tau$ in Eqn. (1.30) results in

$$v_d'(t) = \frac{v_d}{e}$$



time after field is cut off

Hence, the relaxation time τ is the time taken for the drift velocity to decay to 1/e times its initial value.

For metals, the relaxation time $\tau =$ collision time τ_c

From kinetic theory of gases, it is known that

$$\frac{1}{2}m\overline{v}^2 = \frac{3}{2}kT\tag{1.32}$$

Here, \overline{v} is the average velocity of random motion or r.m.s velocity of electrons, k is the Boltzmann's constant in JK⁻¹ and T is the absolute temperature in K.

The average time between two collisions is given by

$$\tau_c = \frac{\lambda}{\overline{\nu}} \tag{1.33}$$

where λ is the mean free path of the electrons in m or the mean distance travelled by the electron between two successive collisions.

Hence,
$$\overline{v} = \frac{\lambda}{\tau_c}$$
 (since $\tau = \tau_c$) (1.34)

From Eqn. (1.32), we have

$$\overline{v} = \sqrt{\frac{3kT}{m}}$$
(1.35)

$$= \sqrt{\frac{3kT}{m}} = \frac{\lambda}{\tau_c} = \frac{\lambda}{\tau} \qquad (\text{as } \tau = \tau_c \text{ for metals})$$
$$\tau = \lambda \sqrt{\frac{m}{3kT}} \qquad (1.36)$$

Therefore,

Substituting Eqn. (1.36) in Eqn. (1.26), we obtain

$$\sigma = \frac{ne^2}{m} \times \lambda \sqrt{\frac{m}{3kT}}$$

$$\sigma = \frac{ne^2 \lambda}{\sqrt{3mkT}}$$
(1.37)

i.e.,

The above equation gives the expression for electrical conductivity.

Hence,
$$\rho = \frac{1}{\sigma} = \frac{\sqrt{3mkT}}{ne^2\lambda}$$
(1.38)

In case of metals, *n* is a constant and λ varies relatively slowly except at very low temperatures. At normal temperatures, λ can be assumed to be a constant. Thus, $\rho \alpha \sqrt{T}$ according to the above theory. But in reality, resistivity is proportional to *T* (experimentally observed). Also, *n* varies from metal to metal. A metal with a larger value of *n* should exhibit higher conductivity which is not the actual case. Hence, there is something basically wrong in the assumptions of classical theory. Despite this, there are some successes of this theory.

EXAMPLE 1.1

The sodium (Na) metal has a BCC structure. The radius of Na atom is 1.85Å. Calculate the electrical resistivity at 0°C, if the classical value of the mean free time at this temperature is 3×10^{-14} s.

Solution Given $\tau = 3 \times 10^{-14}$ s. As Na has BCC structure, the number of atoms per cell, n' = 2 and the radius of Na atom, r = 1.85 Å.

We know that for a BCC structure, $\sqrt{3}a = 4r$

Hence,
$$a = \frac{4r}{\sqrt{3}} = \frac{4}{\sqrt{3}} \times 1.85 \times 10^{-10} = 4.27 \times 10^{-10} \,\mathrm{m}$$

The number of free electrons per unit volume in Na is

$$n = \frac{n'}{a^3} = \frac{2}{(4.27 \times 10^{-10})^3} = 2.57 \times 10^{28} \text{ electrons per } m^3$$

Electrical resistivity,

$$\rho = \frac{m}{ne^2\tau}$$

Here, the mass of an electron, $m = 9.11 \times 10^{-31}$ kg and the charge of an electron, $e = 1.602 \times 10^{-19}$ C.

Therefore,

$$\rho = \frac{9.11 \times 10^{-31}}{2.57 \times 10^{28} \times (1.602 \times 10^{-19})^2 \times (3 \times 10^{-14})}$$
$$= \frac{9.11 \times 10^{-31}}{20.45 \times 10^{-24}} = 4.45 \times 10^{-8} \ \Omega \cdot m$$

EXAMPLE 1.2

A uniform Copper (Cu) wire whose diameter is 0.16 cm carries a steady current of 10 amperes. Its density and atomic weight are 8920 kg/m³ and 63.5 respectively. Calculate the current density and drift velocity of the electrons in Copper (Cu).

Solution Given I = 10 A, d = 0.16 cm $= 0.16 \times 10^{-2}$ m, density of Cu = 8920kg/m³ and atomic weight of Cu = 63.5.

Current density,
$$J = \frac{I}{A} = \frac{I}{\pi \times r^2}$$
, where $r = \frac{d}{2} = \frac{0.16 \times 10^{-2}}{2} = 0.08 \times 10^{-2} \text{ m}$

$$J = \frac{10}{\pi \times (0.08 \times 10^{-2})^2} = \frac{10}{3.14 \times 64 \times 10^{-8}} = 4.98 \times 10^6 \text{ A/m}^2$$

Therefore,

Since $J = nev_d$, the drift velocity is given by

$$v_d = \frac{J}{ne}$$

Electrical Properties of Materials 1.9

Here, *n* is the number of Cu atoms per m^3 as given by

$$n = \frac{n'}{a^3} = \frac{\text{density} \times \text{Avogadro number}}{\text{Atomic weight}} = \frac{8.92 \times 10^3 \times 6.02 \times 10^{26}}{63.5}$$
$$= 8.456 \times 10^{28} / \text{ m}^3$$
$$v_d = \frac{J}{ne} = \frac{4.98 \times 10^6}{(8.456 \times 10^{28}) \times (1.602 \times 10^{-19})} = 3.67 \times 10^{-4} \text{ m/s}$$

Hence,

EXAMPLE 1.3

Calculate the drift velocity of the free electrons with a mobility of 3.5×10^{-3} m²/(V·s) in copper for electric field strength of 0.5 V/m. (AU May 2005)

Solution Given mobility, $\mu = 3.5 \times 10^{-3} \text{ m}^2 / (\text{V} \cdot \text{s})$ and E = 0.5 V/m.

Hence, the drift velocity, $v_d = \mu E = 3.5 \times 10^{-3} \times 0.5 = 1.75 \times 10^{-3} \text{ m/s}$

EXAMPLE 1.4

Resistivity of copper is $1.73 \times 10^{-8} \Omega \cdot m$. Its density is 8.92×10^{3} kg/m³ and atomic weight is 63.5. Assuming classical laws, calculate the mobility of electrons. (AU Dec 2005)

Solution Given resistivity of copper $\rho = 1.73 \times 10^{-8} \Omega \cdot m$, density of Copper = $8.92 \times 10^{3} \text{ kg/m}^{3}$ and atomic weight of Copper = 63.5.

We know that conductivity, $\sigma = ne\mu$

Hence, mobility is given by

$$\mu = \frac{\sigma}{ne} = \frac{1}{\rho ne}$$
 (since $\sigma = 1/\rho$)

Here, the number of charge carriers (electrons) is

$$n = \frac{\text{density} \times \text{Avogadro number}}{\text{Atomic weight}} = \frac{8.92 \times 10^3 \times 6.02 \times 10^{26}}{63.5} = 8.456 \times 10^{28} / \text{m}^3$$

Therefore,

$$\mu = \frac{1}{\rho ne} = \frac{1}{1.73 \times 10^{-8} \times 8.456 \times 10^{28} \times 1.602 \times 10^{-19}} = 0.427 \times 10^{-2} \,\mathrm{m^2} / (\mathrm{V \cdot s})$$

EXAMPLE 1.5

A uniform silver wire has a resistivity of $1.54 \times 10^{-8} \Omega \cdot m$ at room temperature. For an electric field along the wire of 1 volt/cm, compute the average drift velocity of the electrons assuming that there are 5.8×10^{28} conduction electrons per m³. Also, compute the mobility and relaxation time of the electron.

Solution

Given $\rho = 1.54 \times 10^{-8} \Omega \cdot m$, E = 1 V/cm = 100 V/m and $n = 5.8 \times 10^{28} /m^3$.

We know that drift velocity, $v_d = \mu E$

But mobility,

$$\mu = \frac{\sigma}{ne} = \frac{1}{\rho ne} = \frac{1}{1.54 \times 10^{-8} \times 5.8 \times 10^{28} \times 1.602 \times 10^{-19}}$$
$$= 6.99 \times 10^{-3} \text{ m}^2/\text{V} \cdot \text{s}$$

Now,

$$v_d = \mu E = 6.99 \times 10^{-3} \times 100 = 0.699$$
 m/s

Also, the relaxation time is

$$\tau = \frac{\mu m}{e} = \frac{6.99 \times 10^{-3} \times 9.11 \times 10^{-31}}{1.602 \times 10^{-19}} = 3.97 \times 10^{-14} \,\mathrm{s}$$

EXAMPLE 1.6

A certain metal of density 2700kg/m³ and atomic weight 26.98 has a resistivity of $2.62 \times 10^{-8} \Omega \cdot m$ at room temperature. If the applied field is 100 V/m, calculate (i) the mobility of the electrons, (ii) drift velocity of electrons, (iii) the relaxation time and (iv) the mean free path.

Solution Given $\rho = 2.62 \times 10^{-8} \Omega \cdot m$, density = 2700kg/m³, atomic weight = 26.98, *E* = 100 V/m and *T* = 300 K.

(i) Mobility of the electrons,

$$\mu = \frac{\sigma}{ne} = \frac{1}{\rho ne}$$

1

Here,
$$n = \frac{\text{density} \times \text{Avogadro number}}{\text{Atomic weight}} = \frac{2.7 \times 10^3 \times 6.02 \times 10^{26}}{26.98} = 6.02 \times 10^{28} \text{ / m}^3$$

1

Hence.

$$\mu = \frac{1}{\rho ne} = \frac{1}{2.62 \times 10^{-8} \times 6.02 \times 10^{28} \times 1.602 \times 10^{-19}}$$

$$= 3.96 \times 10^{-3} \text{ m}^2/\text{V} \cdot$$

(ii) Drift velocity of electrons, $v_d = uE = 3.96 \times 10^{-3} \times 100 = 0.396$ m/s

(iii) Relaxation time,
$$\tau = \frac{\mu m}{e} = \frac{3.96 \times 10^{-3} \times 9.11 \times 10^{-31}}{1.602 \times 10^{-19}} = 2.25 \times 10^{-14} \text{ s}$$

(iv) Mean free path,
$$\lambda = \tau \sqrt{\frac{3kT}{m}} = 2.25 \times 10^{-14} \times \sqrt{\frac{3 \times 1.38 \times 10^{-23} \times 300}{9.11 \times 10^{-31}}}$$

= 2.25 × 10⁻¹⁴ × 1.17 × 10⁵ = 2.63 × 10⁻⁹ = 2.63 nm

EXAMPLE 1.7

Given that every atom of Cu contributes one electron to the free electron gas and the specific resistance of Cu at 27° C is $4.7 \times 10^{-9} \Omega \cdot m$. The atomic weight of copper is 63.5 and density is 9×10^{3} kg/m³. Compute the mean free path of the electrons.

Solution Given $\rho = 4.7 \times 10^{-9} \Omega \cdot m$, density = $9 \times 10^3 \text{ kg/m}^3$, atomic weight = 63.5 and T = 300 K.

The mean free path, $\lambda = \tau \sqrt{\frac{3kT}{m}}$

Here,

$$\tau = \frac{\mu m}{e} = \frac{m}{\rho n e^2} \qquad \left(\text{since } \mu = \frac{1}{\rho n e} \right)$$

Also,

$$=\frac{\text{density} \times \text{Avogadro number}}{\text{Atomic weight}} = \frac{9 \times 10^3 \times 6.02 \times 10^{26}}{63.5} = 8.53 \times 10^{28} \text{ / m}^3$$

Mean free time,

$$=\frac{m}{\rho ne^2} = \frac{9.11 \times 10^{-91}}{4.7 \times 10^{-9} \times 8.53 \times 10^{28} \times (1.602 \times 10^{-19})^2} = 8.85 \times 10^{-14} \text{ s}$$

0.11 10-31

Now, the mean free path of the electrons is

п

τ

$$\lambda = \tau \sqrt{\frac{3kT}{m}} = 8.85 \times 10^{-14} \times \sqrt{\frac{3 \times 1.38 \times 10^{-23} \times 300}{9.11 \times 10^{-31}}} = 10.35 \times 10^{-9} = 10.3 \text{ nm}$$

EXAMPLE 1.8

The resistivity of Copper at 20°C is $1.69 \times 10^{-8} \Omega \cdot m$ and density of free electrons, $n = 8.5 \times 10^{28}/m^3$. Calculate the mean free path. Given at 20°C, $\overline{c} = 1.154 \times 10^5$ m/s.

Solution Given $\rho = 1.69 \times 10^{-8} \Omega \cdot m$, density = $8.5 \times 10^{28}/m^3$ and $\overline{c} = 1.154 \times 10^5 m/s$.

The mean free time, $\tau = \frac{m}{\rho n e^2} = \frac{9.11 \times 10^{-31}}{1.69 \times 10^{-8} \times 8.5 \times 10^{28} \times (1.602 \times 10^{-19})^2} = 2.47 \times 10^{-14} \text{ s}$

Now, the mean free path of the electrons, $\lambda = \tau \overline{c} = 2.47 \times 10^{-14} \times 1.154 \times 10^5 = 2.85$ nm

EXAMPLE 1.9

The density and atomic weight of metallic copper are 8.93×10^3 kg/m³ and 63.5 kg respectively. If the electrical conductivity of copper is 6×10^7 ($\Omega \cdot m$)⁻¹, calculate the mean free time. The mass of a free electron = 9.1×10^{-31} kg. (AU Nov 2003)

Solution Given density of copper = 8.93×10^3 kg/m³, atomic weight of copper = 63.5 kg, electrical conductivity of copper, $\sigma = 6 \times 10^7 (\Omega \cdot m)^{-1}$ and mass of a free electron, $m = 9.11 \times 10^{-31}$ kg.

Number of conduction electrons per m^3 is

 $n = \frac{\text{Density} \times \text{Number of free electrons per atom} \times \text{Avagadro number}}{\text{Atomic weight}}$ $= \frac{8.93 \times 10^3 \times 1 \times 6.023 \times 10^{26}}{63.5} = 8.47 \times 10^{28} / \text{m}^3$

The electrical conductivity, $\sigma = \frac{ne^2\tau}{m}$

Therefore, mean free time,
$$\tau = \frac{m\sigma}{ne^2} = \frac{9.1 \times 10^{-31} \times 6 \times 10^7}{8.47 \times 10^{28} \times (1.6 \times 10^{-19})^2} = 2.518 \times 10^{-14} \text{ s}$$

EXAMPLE 1.10

The density and atomic weight of copper are 8900 kg/m³ and 63.5. The relaxation time of electron in Copper at 300 K is 10⁻¹⁴s. Assuming that each copper atom contributes one free electron, calculate the electrical conductivity of copper. (AU June 2009)

Solution Given density of copper = 8.9×10^3 kg/m³, atomic weight of copper = 63.5 and relaxation time = 10^{-14} s.

Number of conduction electrons per m^3 is

$$n = \frac{\text{Density} \times \text{Number of free electrons per atom} \times \text{Avagadro number}}{\text{Atomic weight}}$$
$$= \frac{8.9 \times 10^3 \times 1 \times 6.023 \times 10^{26}}{63.5} = 8.44 \times 10^{28} / \text{m}^3$$
$$\text{y, } \sigma = \frac{ne^2\tau}{63.5} = \frac{8.44 \times 10^{28} \times (1.602 \times 10^{-19})^2 \times 10^{-14}}{63.5}$$

Electrical conductivity, $\sigma = \frac{ne^2\tau}{m} = \frac{8.44 \times 10^{28} \times (1.602 \times 10^{-19})^2 \times 10^{-14}}{9.11 \times 10^{-31}}$

$$=\frac{2.16064\times10^{-23}}{9.11\times10^{-31}}=2.3717\times10^{7}(\Omega\cdot\mathrm{m})^{-1}$$

EXAMPLE 1.11

A copper wire has a resistivity of $1.7 \times 10^{-8} \Omega \cdot m$ at room temperature of 300K. If the copper is highly pure, find the resistivity at 700°C.

Solution Given $\rho = 1.7 \times 10^{-8} \Omega \cdot m$.

We know that $\rho \alpha \sqrt{T}$

i.e.,

$$\frac{\rho_1}{\rho_2} = \sqrt{\frac{T_1}{T_2}}$$

(or)
$$\rho_2 = \rho_1 \sqrt{\frac{T_2}{T_1}} = 1.7 \times 10^{-8} \times \sqrt{\frac{700 + 273}{300}} = 3.06 \times 10^{-8} \ \Omega \cdot m$$

1.5 THERMAL CONDUCTIVITY

Thermal conduction is the process by which heat energy is transmitted from the hotter to colder part of a body without the motion of the particles of the body. Ex. heating a metal rod at one end.

The coefficient of thermal conductivity K is defined as the amount of heat flowing through unit area of cross section of the metal normal to it in one second under unit temperature gradient.

Consider heat flowing from the hot end of a metal rod at temperature T_1 K to cold end at temperature T_2 K as shown in Fig. 1.5. Here, x is the distance between the hot and cold end in m, A is the area of cross section in m² and heat flow is normal to the area of cross section.

The quantity of heat Q flowing or conducted along the rod for t seconds can be written as follows

$$Q \propto A$$

$$Q \propto (T_1 - T_2) \text{ (Temp. difference)}$$

$$Q \propto t \text{ (time of heat flow)}$$

$$Q \propto \frac{1}{x}$$

$$T_1 \qquad T_2 \longrightarrow \text{Heat flow}$$

$$\text{Hot end} \qquad \text{Cold end}$$

$$| \leftarrow x \longrightarrow |$$

$$A \text{ (Area of cross section)}$$

Fig. 1.5 Heat flow in metal rod

Combining all the above equations, we get

$$Q \propto \frac{A(T_1 - T_2)t}{x}$$
$$Q = \frac{KA(T_1 - T_2)t}{x}$$

i.e.,

where *K* is the coefficient of thermal conductivity. Hence,

$$\frac{Q/t}{A} = \frac{K(T_1 - T_2)}{x}$$

$$K = \frac{\left(\frac{Q/t}{A}\right)}{\frac{(T_1 - T_2)}{x}}$$
(1.39)

i.e.,

where $\frac{(T_1 - T_2)}{x}$ is the temperature gradient.

If $A = 1 \text{ m}^2$, t = 1 sec, $(T_1 - T_2) = 1 \text{ K}$ and x = 1 m, then Q = K which leads us to the definition of thermal conductivity as stated in the beginning. The temperature gradient $(T_1 - T_2)/x$ can be written as -dT/dx. Here, negative sign indicates that temperature decreases as the distance from the hot end increases.

Here, the unit of coefficient of thermal conductivity K is given by

$$K = \frac{\left(\frac{Q/t}{A}\right)}{\frac{(T_1 - T_2)}{x}} = \frac{W/m^2}{K/m} = \frac{W}{m \cdot K}$$

1.5.1 Expression for Thermal Conductivity

Consider a plane CD in a conductor normal to the direction of flow of heat as shown in Fig. 1.6. in which two other planes AB and EF on either sider of CD are at a distance equal to the mean free path of the electron λ . In solids, heat conduction is due to both electrons and phonons. Phonons are quanta of mechanical energy due to lattice vibrations.

Hence,

$$K_{\text{Total}} = K_{e^-} + K_{\text{Phonon}}$$

In metals, contribution due to electrons is much more than phonons.

Therefore,

and hence,

 $K_{e^-} >> K_{\text{Phonon}}$ $K_{\text{Total}} = K_{e^-}$



Fig. 1.6 Plane CD in a conductor normal to the direction of flow of heat

Let the temperatures of planes AB and EF be T_1 K and T_2 K respectively with $T_1 > T_2$ and the distance between AB and EF be 2λ . The kinetic energy of electrons at the hotter end is greater than that at colder end. As a result of collisions, electrons at hot end lose kinetic energy while at cold end gain energy. Thus, the conduction of heat through the element is due to transfer of energy.

From kinetic theory of gases as shown in Fig. 1.7, if n = number of electrons per m³ and \overline{v} is the mean velocity of each electron while crossing CD, then the number of electrons crossing CD per unit area per second is

given by
$$\frac{1}{6}n\overline{v}$$
.

Hence, heat transferred from left to right per second across unit area of CD

$$= \left(\frac{1}{6}n\overline{v}\right)\frac{3}{2}kT_1 \text{ (since } \frac{1}{2}m\overline{v_1}^2 = \frac{3}{2}kT_1 \text{ where } \overline{v_1} \text{ is the velocity of electrons}$$

 $-x \leftrightarrow x$ $-y \leftrightarrow -z$ Fig. 1.7 Kinetic theory of gases

corresponding to temperature T_1 and k is the Boltzmann's constant).

Similarly, heat transferred from right to left per second across unit area of $CD = \left(\frac{1}{6}n\overline{v}\right)\frac{3}{2}kT_2$ (since $\frac{1}{2}m\overline{v_2}^2 = \frac{3}{2}kT_2$ where $\overline{v_2}$ is the velocity of electrons corresponding to temperature T_2).

Here, the net energy transferred from left to right /unit area/sec is the difference between the heat transferred from left to right and right to left. Therefore, the amount of heat flowing through unit area of cross section is

$$\frac{Q/t}{A} = \left(\frac{1}{6}n\overline{v}\right)\frac{3}{2}k(T_1 - T_2)$$
$$= \frac{n\overline{v}k}{4}(T_1 - T_2)$$
(1.40)

and the temperature gradient is $\frac{(T_1 - T_2)}{2\lambda}$.

(1.41)

The coefficient of thermal conductivity K is the amount of heat flowing through unit area of cross section of the metal normal to it in one second under unit temperature gradient.

i.e.,
$$K = \frac{\text{amount of heat flowing through unit area of cross section}}{\text{temperature gradient}} = \frac{\left(\frac{Q/t}{A}\right)}{\frac{(T_1 - T_2)}{2\lambda}}$$

Therefore,

$$K = \frac{\frac{nvk(T_1 - T_2)}{4}}{\frac{(T_1 - T_2)}{2\lambda}}$$
$$K = \frac{n\overline{v}k\lambda}{2}$$

i.e.,

The above equation provides the expression for thermal conductivity.

1.6 WIEDEMANN-FRANZ LAW AND LORENTZ NUMBER

Wiedemann-Franz Law states that the ratio of thermal conductivity K to the electrical conductivity σ in a metal is directly proportional to the absolute temperature of the metal.

i.e.,

$$\frac{K}{\sigma} \propto T$$

or

$$\frac{K}{\sigma} = LT$$

i.e.,

$$\frac{K}{\sigma T} = L \text{ (a constant)} \tag{1.42}$$

Here, L is called the Lorentz number. The value of L is $2.44 \times 10^{-8} \text{ W} \cdot \Omega \cdot \text{K}^{-2}$ at 293 K (experimental value).

1.6.1 Derivation of Wiedemann-Franz Law

The thermal conductivity is given by

$$K = \frac{n\overline{\nu}k\lambda}{2} \tag{1.43}$$

and the electrical conductivity is given by

$$\sigma = \frac{ne^2\tau}{m} = \frac{ne^2}{m} \left(\frac{\lambda}{\overline{\nu}}\right) \qquad (\text{since } \tau = \lambda / \overline{\nu}) \qquad (1.44)$$

Dividing Eqn. (1.43) by Eqn. (1.44), we get

$$\frac{K}{\sigma} = \frac{n\overline{v}k\lambda}{2} \times \frac{m\overline{v}}{ne^2\lambda}$$
$$\frac{K}{\sigma} = \frac{1}{2}km\left(\frac{\overline{v}}{e}\right)^2$$
(1.45)

i.e.,

From kinetic theory of gases, it follows that

$$\frac{1}{2}m\overline{v}^2 = \frac{3}{2}kT$$
(1.46)

Substituting Eqn. (1.46) in Eqn. (1.45), we get

$$\frac{K}{\sigma} = \frac{3}{2}kT \cdot \frac{k}{e^2} = \left(\frac{3}{2}T\right) \left(\frac{k}{e}\right)^2$$
$$\frac{K}{\sigma T} = \frac{3}{2} \left(\frac{k}{e}\right)^2 = L, \text{ the Lorentz number}$$
(1.47)

or

In other words,

$$\frac{K}{L} = L$$
, which is a constant (1.48)

 σT Substituting the values for '*b*' the Boltzmann's constant and *a*, the charge of the electron i

Substituting the values for 'k' the Boltzmann's constant and *e*, the charge of the electron in Eqn. (1.47), we obtain the Lorentz number as $L = 1.12 \times 10^{-8} \text{ W} \cdot \Omega \cdot \text{K}^{-2}$.

Let us check this by substituting the experimental values of *K* and σ for a metal say copper, Cu. For Cu at 293 K (20°C), the electrical resistivity, $\rho = \frac{1}{\sigma} = 1.72 \times 10^{-8} \Omega \cdot m$ and the thermal conductivity, $K = 386 \text{ W/m} \cdot \text{K}$.

The Lorentz number is

$$L = \frac{K}{\sigma T} = \frac{K\rho}{T} = \frac{386 \times 1.72 \times 10^{-8}}{293} = 2.266 \times 10^{-8} \,\mathrm{W} \cdot \Omega \cdot \mathrm{K}^{-2}$$

This value of L does not agree with the value calculated from classical formula. The reason for this discrepancy lies in the wrong assumption that all the free electrons take part in the conduction process.

According to quantum free electron theory, the electrons near the Fermi level alone contribute towards thermal and electrical conduction. Using the quantum free electron model, it can be proved that,

$$\frac{K}{\sigma} = \frac{\pi^2}{3} \left(\frac{k}{e}\right)^2 T$$
$$\frac{K}{\sigma T} = \frac{\pi^2}{3} \left(\frac{k}{e}\right)^2 = L$$

or

which yields the correct experimental value of Lorentz number L.

1.7 SUCCESS AND FAILURES OF CLASSICAL FREE ELECTRON THEORY

The successes and failures of classical free electron theory are given as follows:

Successes

- (i) Verifies Ohm's law.
- (ii) Explains electrical and thermal conductivities of metals.
- (iii) Verifies that $K/\sigma T$ is a constant (Wiedemann-Franz Law).
- (iv) Explains optical properties of materials.

Failures

- (i) $\rho \propto \sqrt{T}$ from classical theory but experimentally $\rho \propto T$.
- (ii) The value of specific heat from this theory is 4.5 R where R is the Universal gas constant, but the experimental value is 3R. From classical theory, electronic specific heat is 3R/2. But the actual value is 0.01 R only.
- (iii) Conductivity of semiconductors and insulators is not explained.
- (iv) The value of $K/\sigma T$ at room temperature is a constant but the value of this constant deviates from the excremental value. Further, at low temperatures, $K/\sigma T$ is not a constant.
- (v) Ferromagnetism and paramagnetism are not explained.
- (vi) Compton effect, photoelectric effect and black body radiation are not explained.

EXAMPLE 1.12

The copper has electrical conductivity at 300 K as 6.4×10^7 mho/m. Calculate the thermal conductivity of copper and the Lorentz number is $L = 2.45 \times 10^{-8}$ W $\cdot \Omega \cdot K^{-2}$.

Solution Given T = 300 K, $\sigma = 6.4 \times 10^7 \text{ mho/m}$ and $L = 2.45 \times 10^{-8} \text{ W} \cdot \Omega \cdot \text{K}^{-2}$.

The Lorentz number, $L = \frac{K}{\sigma T}$

Therefore, the thermal conductivity (K) is

$$K = L\sigma T = 2.45 \times 10^{-8} \times 6.4 \times 10^{7} \times 300 = 470.4 \text{ W/m} \cdot \text{K}$$

EXAMPLE 1.13

The thermal conductivity of Zinc (Zn) is 327 W/m·K at 300 K. Compute the Lorentz number if the density of Zn is 7.13×10^3 kg/m³ and atomic weight is 63.38. The mean relaxation time in the divalent Zn is 2.5×10^{-14} s.

Solution Given K = 327 W/m·K, T = 300 K, density $= 7.13 \times 10^3$ kg/m³, atomic weight = 63.38 and $\tau = 2.5 \times 10^{-14}$ s.

For divalent Zinc, the number of free electrons / m³ is

$$n = 2 \times \frac{\text{density} \times \text{Avogadro number}}{\text{Atomic weight}} = 2 \times \frac{7.13 \times 10^3 \times 6.02 \times 10^{20}}{63.38}$$
$$= 1.35 \times 10^{29} / \text{m}^3$$

The electrical conductivity,

$$\sigma = \frac{ne^2\tau}{m} = \frac{1.35 \times 10^{29} \times (1.602 \times 10^{-19})^2 \times 2.5 \times 10^{-14}}{9.11 \times 10^{-31}}$$

$$= 0.95 \times 10^8 (\Omega \cdot m)^{-1}$$
 or $0.95 \times 10^{\circ}$ mho/m

Lorentz number, $L = \frac{K}{\sigma T} = \frac{327}{0.95 \times 10^8 \times 300} = 1.15 \times 10^{-8} \text{ W} \cdot \Omega \cdot \text{K}^{-2}$

EXAMPLE 1.14

The thermal conductivity of a metal is 123.92 W/m·K. Find the electrical conductivity and Lorentz number when the metal possesses a relaxation time of 10^{-14} seconds at 300 K and the density of electrons is $6 \times 10^{28} / \text{m}^3$.

Solution Given K = 123.92 W/m·K, $\tau = 10^{-14}$ s and the density of electrons, $n = 6 \times 10^{28}$ /m³.

We know that the electrical conductivity,

$$\sigma = \frac{ne^2\tau}{m} = \frac{6 \times 10^{28} \times (1.602 \times 10^{-19})^2 \times 10^{-14}}{9.11 \times 10^{-31}}$$
$$= 1.69 \times 107 \ (\Omega \cdot m)^{-1} \text{ or } 1.69 \times 10^7 \text{ mho/m}$$

$$= 1.69 \times 107 \ (\Omega \cdot m)^{-1} \text{ or } 1.69 \times 10^{7} \text{ mho}$$

Lorentz number,

$$L = \frac{K}{\sigma T} = \frac{123.92}{1.69 \times 10^7 \times 300} = 2.44 \times 10^{-8} \text{ W} \cdot \Omega \cdot \text{K}^{-2}$$

EXAMPLE 1.15

The thermal conductivity of copper at 300 K is 470.4 W/m·K. Calculate the electrical conductivity of copper at 300 K. Lorentz number $L = 2.45 \times 10^{-8}$. (AU June 2010)

Solution Given K = 470.4. W/m \cdot K at T = 300 K.

From Wiedemann - Franz Law,

$$\frac{K}{\sigma} = LT$$

Hence, electrical conductivity of copper at 300 K is

$$\sigma = \frac{K}{LT} = \frac{470.4}{2.45 \times 10^{-8} \times 300} = \frac{470.4}{7.35 \times 10^{-6}}$$
$$= 6.4 \times 10^{7} (\Omega \, \text{m})^{-1} \text{ or } 6.4 \times 10^{7} \text{ mho/m}$$

1.8 QUANTUM FREE ELECTRON THEORY

It is known that the classical free electron theory does not explain the variation of resistivity with temperature, value of Lorentz number and the properties of conductors, semiconductors and insulators as well as black body radiation, photoelectric effect etc.

The Planck's Quantum theory correlates theory and experimental observations. Also, Einstein published a paper using Planck's idea of quanta for explaining the photoelectric effect. He wrote: "In accordance with the assumption to be considered here, the energy of a light ray spreading out from a point source is not continuously distributed over an increasing space, but consists of a finite number of energy quanta which are localized at points in space, which move without dividing, and which can only be produced and absorbed as complete units". In other words, Einstein proposed that *light behaved not like a wave, but like a particle: the photon*.

Sommerfield and Pauli extended the quantum concept to the free electron theory. Using a combination of quantum concept and the Louis de-Broglie wave nature of matter, Schrödinger derived wave equations which yielded a correct explanation for questions unanswered by classical theory like variation of resistivity with temperature, specific heat etc.,

The Quantum free electron theory is discussed in detail. Sommerfield and Pauli preserved some features of the classical theory such as

- (i) Electrons travel under constant potential but confine to the boundaries of metal.
- (ii) Mutual repulsion between electrons is neglected.
- (iii) The attraction between electrons and the positive ion cores is ignored.

Additional features introduced by them as per Quantum free electron theory are

- (i) The wave nature of electrons is taken into account.
- (ii) The energies of free electrons are quantized
- (iii) The distribution of electrons is as per the Pauli's exclusion principle.
- (iv) The electrons obey Fermi Dirac statistics

1.9 ELECTRONS IN A METAL-PARTICLE IN A THREE-DIMENSIONAL BOX

The real picture of a free electron trapped in a metal can be compared to the particle in a three-dimensional box problem which is the application of Schrödinger's equation in 3D to a free particle. (Assumptions No. (i) to (iv) of additional features as per Quantum free electron theory).

Consider the case of a single particle i.e., an electron of mass *m*, confined within a rectangular box with edges parallel to *X*, *Y* and *Z* axes as shown in Fig. 1.8. Let the sides of rectangular box be *a*, *b* and *c* respectively. The particle can move freely within the region 0 < x < a, 0 < y < b and 0 < y < c i.e., inside the box where potential *V* is zero.

From Fig. 1.8, it is seen that for a particle in a box, the potential is

$$V(x, y, z) = 0, 0 < x < a$$

$$V(x, y, z) = 0, 0 < x < b$$

$$V(x, y, z) = 0, 0 < x < c$$



Fig. 1.8 An electron of mass m, confined within a rectangular box

and

The potential rises suddenly to have a very large value at the boundaries i.e., the potential outside the box is infinite. The Schrödinger wave equation inside the box is given by

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{2m}{\hbar^2} E\psi = 0$$
(1.49)

This is a partial differential equation in three independent variables and may be solved by the method of separation of variables. The solution of Eqn. (1.49) is of the following form.

$$\Psi(x, y, z) = X(x)Y(y)Z(z) = XYZ$$
(1.50)

where X(x) is a function of x alone, Y(y) is a function of y alone and Z(z) is a function of z alone.

Substituting the value of ψ from Eqn. (1.50) in Eqn. (1.49) and dividing by X(x)Y(y)Z(z), we get

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} + \frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} + \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} + \frac{2mE}{\hbar^2} = 0$$

The above equation may be written as

 $\frac{1}{X}\frac{\partial^2 X}{\partial x^2} = k_x$

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} = -\frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} - \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} - \frac{2mE}{\hbar^2}$$
(1.51)

The left hand side of Eqn. (1.51) is a function of x alone, while the right hand is a function of y and z and is independent of x. Both sides are equal to each other. This is only possible when they are separately equal to a constant quantity.

i.e.,

$$-\frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} - \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} - \frac{2mE}{\hbar^2} = k_x$$

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} = -\frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} - \frac{2mE}{\hbar^2} - k_x$$
(1.53)

(1.52)

or

The left hand side of Eqn. (1.53) is a function of y alone while right hand side is a function of z and is independent of y. If above equations are to be satisfied, both sides must be equal to a constant, say, k_y

i.e.,
$$\frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} = k_y \tag{1.54}$$

and

$$-\frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} - \frac{2mE}{\hbar^2} - k_x = k_y$$

or

$$\frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} = -\frac{2mE}{\hbar^2} - k_x - k_y \tag{1.55}$$

Again, we have

$$\frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} = k_z$$

Electrical Properties of Materials 1.21

$$\frac{2mE}{\hbar^2} - k_x - k_y = k_z \tag{1.56}$$

or

$$\frac{2mE}{\hbar^2} = k_x + k_y + k_z \tag{1.57}$$

For convenience, we introduce

$$k_x = -\frac{2mE_x}{\hbar^2}, \ k_y = -\frac{2mE_y}{\hbar^2} \text{ and } k_z = -\frac{2mE_z}{\hbar^2}$$

Now the differential equation in x, y and z coordinates may be written as

$$\frac{\partial^2 X}{\partial x^2} + \frac{2mE_x}{\hbar^2} X = 0 \tag{1.58a}$$

$$\frac{\partial^2 Y}{\partial y^2} + \frac{2mE_y}{\hbar^2} Y = 0 \tag{1.58b}$$

and

$$\frac{\partial^2 Z}{\partial z^2} + \frac{2mE_z}{\hbar^2} Z = 0 \tag{1.58c}$$

The general solution of Eqn. (1.58a) will be a sine function of arbitrary amplitude, frequency and phase i.e.,

$$X(x) = A\sin(Bx + C) \tag{1.59}$$

where A, B and C are constants whose values are determined by boundary conditions.

Here, $|\psi|^2$ represents the probability of finding the particle at any point within the box. Therefore, $|X(x)|^2$ which is a function of *x* coordinates only, represents the probability of finding the particle at any point along the *X*-axis. As the potential is very high at the walls of the box, the probability of finding the particle at the walls will be zero.

i.e.,	$\left X(x)\right ^2 = 0$	when $x = 0$ and $x = a$
or	X(x) = 0	when $x = 0$ and $x = a$

Using these boundary conditions in Eqn. (1.59), we have

$$0 = A\sin\left(0 + C\right)$$

and as $A \neq 0$, sin C = 0 which yields C = 0

Similarly, $0 = A \sin Ba$ which yields $\sin Ba = 0$ or $Ba = n_x \pi$

i.e.,
$$B = \frac{n_x \pi}{a}$$

where n_x is a positive integer.

Hence,
$$X(x) = A \sin \frac{n_x \pi x}{a}$$
 (1.60)

Applying the normalization condition between x = 0 to x = a, we have

$$\int_{0}^{a} \left| X(x) \right|^2 dx = 1$$

$$\int_{0}^{a} \left| A \sin \frac{n_{x} \pi x}{a} \right|^{2} dx = 1$$

$$A^{2} \int_{0}^{a} \sin^{2} \frac{n_{x} \pi x}{a} dx = 1$$

$$A^{2} \frac{a}{2} = 1 \text{ or } A = \sqrt{\left(\frac{2}{a}\right)}$$

$$X(a) = \sqrt{\left(\frac{2}{a}\right)} \sin \frac{n_{x} \pi x}{a}$$
(1.61)

Therefore,

 $X(a) = \sqrt{\left(\frac{-}{a}\right)\sin\frac{x}{a}}$ (1.61)

Similarly, upon solving Eqns. (1.58b) and (1.58c), we obtain

 $\overline{2}$

$$Y(y) = \sqrt{\frac{2}{b}} \sin \frac{n_y \pi y}{b}$$
(1.62)

(1.63)

and

$$Z(z) = \sqrt{\frac{2}{c}} \sin \frac{n_z \pi z}{c}$$

The complete wave function

$$\begin{split} \psi_{n_x,n_y,n_z}(x,y,z) &\text{ has the form} \\ \psi_{n_x,n_y,n_z}(x,y,z) &= X(x)Y(y)Z(z) \\ &= \sqrt{\left(\frac{2}{a}\right)} \sin \frac{n_x \pi x}{a} \sqrt{\left(\frac{2}{b}\right)} \sin \frac{n_y \pi y}{b} \sqrt{\left(\frac{2}{b}\right)} \sin \frac{n_z \pi z}{c} \\ &= \frac{2\sqrt{2}}{\sqrt{abc}} \sin \left(\frac{n_x \pi x}{a}\right) \sin \left(\frac{n_y \pi y}{a}\right) \sin \left(\frac{n_z \pi z}{a}\right) \end{split}$$
(1.64)

From Eqn.(1.61), we have

$$\frac{\partial^2 X}{\partial x^2} = -\left(\frac{n_x \pi x}{a}\right)^2 \sqrt{\frac{2}{a}} \sin \frac{n_x \pi x}{a} = -\left(\frac{n_x \pi}{a}\right)^2 X(x)$$
(1.65)

Substituting the value from Eqn. (1.65) in Eqn. (1.58a), we have

$$-\left(\frac{n_x \pi x}{a}\right)^2 X(x) + \frac{2m}{\hbar^2} E_x X(x) = 0$$
$$E_x = \frac{1}{2m} \left(\frac{n_x \pi \hbar}{a}\right)^2 = \frac{n_x^2 h^2}{8ma^2} \qquad (\text{since } \hbar = \frac{h}{2\pi})$$
(1.66)

or

Electrical Properties of Materials 1.23

Similarly,
$$E_y = \frac{n_y^2 h^2}{8 m b^2}$$
(1.67)

and

$$E_z = \frac{n_z^2 h^2}{8 mc^2}$$
(1.68)

The allowed values of total energy are given by

$$E = E_x + E_y + E_z$$

$$= \frac{h^2}{8m} \left[\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right]$$
(1.69)

where n_x , n_y and n_z denote any set of three positive numbers.

When the box is a cube, i.e., a = b = c, the energy expression is given by

$$E_n = \frac{h^2}{8ma^2} \left[n_x^2 + n_y^2 + n_z^2 \right] \text{ with } n_x, n_y, n_z = 1, 2, 3, \dots$$
(1.70)

and the wave function is given by

$$\Psi_n = \sqrt{\frac{8}{a^3}} \sin\left(\frac{n_x \pi x}{a}\right) \sin\left(\frac{n_y \pi y}{a}\right) \sin\left(\frac{n_z \pi z}{a}\right)$$
(1.71)
$$n^2 = n_x^2 + n_y^2 + n_z^2$$

Here,

Figure 1.9 shows the energy levels for few states of a particle enclosed in a cubical box.



Fig. 1.9 Energy levels for few states of a particle enclosed in a cubical box

From Fig. 1.9, it is clear that

- (i) Three integers n_x , n_y and n_z called quantum numbers are required to completely specify an energy state. Since the particle exists in the box, none of the quantum numbers can be zero. The minimum value would be (1,1,1).
- (ii) *E* depends on the sum of squares of the quantum numbers and not on the individual values.

(iii) Several combination of the three quantum numbers may give rise to different wave functions but same energy value. Such states are called *degenerate states*.

Degeneracy

Different wave functions with three different quantum numbers may lead to same value of energy. This is referred to as *degeneracy*. Ex: Three independent states having quantum numbers (112); (121) and (211) have the same energy as shown below

$$E_{211,112,121} = \frac{6h^2}{8ma^2}$$

These levels are triply degenerated or 3 fold degenerate. The levels (111) (222) etc., are non-degenerate. Thus ψ_{111} , ψ_{222} etc. are non degenerate and ψ_{211} , ψ_{121} , ψ_{112} are degenerate. Degeneracy can be broken by application of a magnetic field or electric field to the system.

Electrons in a metal

Consider a metal piece of macroscopic dimension i.e., 1 cm³ (i.e., a 3D box with side 1 cm)

For ground state energy, $n_x = n_y = n_z = 1$

i.e.,

$$E_{111} = \frac{h^2}{8ma^2} (1^2 + 1^2 + 1^2)$$
$$E_{111} = \frac{3h^2}{8ma^2}$$

or

or

Substituting the values of *h*, *m* and *a*, we have

 $E_{111} = 1.806 \times 10^{-33} \text{ J}$ $E_{111} = 1.127 \times 10^{-14} \text{ eV}$ (since 1 eV = 1.602 × 10⁻¹⁹ J)

To calculate separation between consecutive energy levels, consider the next level after ground state i.e., (211).

$$E_{211} = \frac{h^2}{8 m a^2} (2^2 + 1^2 + 1^2) = \frac{6h^2}{8 m a^2}$$

$$E_{111} = \frac{2}{8} \frac{2}{10} (2^2 + 1^2 + 1^2) = \frac{6h^2}{8 m a^2}$$

i.e., $E_{211} = 2.258 \times 10^{-5}$

The separation between consecutive levels is of the order of 10^{-14} eV. The energy of the ground state is around 10^{-14} eV and can be considered zero. As consecutive levels are separated by a small value, the energy distribution can be regarded as *quasi-continuous*. The above results indicate that even the pool of free electrons (electron gas) have quantised energies and thus have a distribution of energies. They do not possess the same average energy of the order of kT as indicated in classical theory.

In order to find the distribution of energies, it is necessary to calculate the *density of states* i.e., the number of states available per eV per m^3 for electron occupancy and the probability that a quantum state of energy *E* is occupied by an electron.

EXAMPLE 1.16

An electron in a cubical box of metal is subjected to the influence of a magnetic field such that its length 'L' increases by dL with its breath 'b' and thickness 'd' remains the same. Show that the three fold degeneracy of the 311 state now reduces to two fold.

Solution Given that $E_{311} = E_{113} = E_{131}$ before the application of magnetic field.

The total energy of an electron in a cubical box is

$$E_{n_x n_y n_z} = \frac{h^2}{8 m} \left[\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right]$$

 $E_{311} = \frac{h^2}{8m} \left[\frac{9+1+1}{L^2} \right]$

Here,

$$E_{311} = \frac{11h^2}{8mL^2} \quad \text{(since the box is cubical)}$$
$$E_{111} = \frac{3h^2}{8mL^2}$$

Therefore, $\Delta E = E_{311} - E_{111} = \frac{8h^2}{8mL^2} = \frac{h^2}{mL^2}$ (before the field is applied)

Also, length L increases on application of magnetic field for 311, i.e., L = L + dL.

Hence,

$$E_{311} = \frac{h^2}{8m} \left[\frac{9}{(L+dL)^2} + \frac{1^2}{L^2} + \frac{1^2}{L^2} \right] = \frac{h^2}{8m} \left[\frac{9}{(L+dL)^2} + \frac{2}{L^2} \right]$$

and

 $E_{111} = \frac{h^2}{8m} \left[\frac{1}{(L+dL)^2} + \frac{2}{L^2} \right]$ $\Delta E = E_{311} - E_{111}$

Therefore,

$$= \frac{h^2}{8m} \left[\frac{9}{(L+dL)^2} + \frac{2}{L^2} - \frac{1}{(L+dL)^2} - \frac{2}{L^2} \right]$$
$$= \frac{8h^2}{(L+dL)^2 8m} = \frac{h^2}{m(L+dL)^2} \qquad (after the$$

(after the field is applied)

Similarly, for E_{113} and E_{131} , we have

$$E_{113} = E_{131} = \frac{h^2}{8m} \left[\frac{1}{\left(L + dL\right)^2} + \frac{1}{L^2} + \frac{3^2}{L^2} \right]$$

$$= \frac{h^2}{8m} \left[\frac{1}{(L+dL)^2} + \frac{10}{L^2} \right]$$

 $\Delta E = E_{113} - E_{111}$

Also,

$$= \frac{h^2}{8m} \left[\frac{1}{(L+dL)^2} + \frac{10}{L^2} - \frac{1}{(L+dL)^2} - \frac{2}{L^2} \right]$$
$$= \frac{8h^2}{8mL^2} = \frac{h^2}{mL^2}$$

For E_{113} and E_{131} , $\Delta E = \frac{h^2}{mL^2}$, which is same as before and after the application of field.

For
$$E_{311}$$
, $\Delta E_{311} = \frac{h^2}{m(L+dL)^2}$

where L = 1 nm and dL = 0.5 nm.

Therefore,
$$\Delta E_{311} = \frac{(6.627 \times 10^{-34})^2}{9.11 \times 10^{-31} \times (1.5 \times 10^{-9})^2} = 2.145 \times 10^{-19} \text{ J}$$

$$=\frac{2.145\times10^{-19}}{1.602\times10^{-19}}=1.34 \text{ eV } 20.50$$

EXAMPLE 1.17

An electron is confined to one dimensional infinite potential well of width 0.2 nm. It is found that when the energy of the particle is 230 eV, its eigen function has 5 antinodes. Find the mass of the particle and show that it can never have an energy equal to 1 keV.

Solution When the eigen function has 5 antinodes, it represents the fifth energy level i.e., n = 5. For this case,

$$E_5 = n^2 E_1 = 230 \text{ eV}$$

Therefore,

$$E_1 = \frac{E_5}{n^2} = \frac{230 \times 1.602 \times 10^{-19}}{5^2} \qquad \text{(since 1 eV} = 1.602 \times 10^{-19} \text{ J}\text{)}$$
$$= 14.74 \times 10^{-19} \text{ J}$$
$$h^2$$

We know that, $E_1 = \frac{h^2}{8ma^2}$

where Planck's constant, $h = 6.627 \times 10^{-34} \text{ J} \cdot \text{s}$

Hence,
$$m = \frac{h^2}{8E_1a^2} = \frac{(6.627 \times 10^{-34})^2}{8 \times 14.7 \times 10^{-19} \times (0.2 \times 10^{-9})^2} = 9.34 \times 10^{-31} \text{ kg}$$
For $E_n = 1 \text{ keV} = 10^3 \text{ eV}$,

$$n = \sqrt{\frac{E_n}{E_1}} = \sqrt{\frac{1 \times 10^3 \times 1.602 \times 10^{-19}}{14.74 \times 10^{-19}}} = 10.425$$

As *n* is not an integer, $E_n = 1$ keV is not a permitted value of energy. So, the particle can never have an energy equal to 1 keV.

1.10 DENSITY OF ENERGY STATES

In order to calculate the density of energy states, the potential energy field in a metal and the motion of electrons inside a metal are to be discussed. Consider the potential energy distribution within and at the surface of the metal. (i.e., for a 3D array of atoms), the resultant potential energy at any point in the metal is simply the sum of potential energies produced at this point by all the ions in the crystal lattice. An atom of atomic number z has a net positive charge Z_e on its nucleus.

Surrounding it we have the electron cloud with Z electrons. By Gauss's law, the potential any point at a distance r from the nucleus varies inversely as r and directly as the total charge enclosed within a sphere of radius r. The potential energy is given by, U = -eV.

The potential at any point may be chosen as the zero reference of potential because only differences in potential have physical significance and hence, we choose zero potential at infinity. Then the potential at any point would be negative. The potential energy curves can then be drawn as shown in Fig. 1.10, where α represents the nucleus.

The vertical scale gives the potential energy, U, the horizontal scale gives the radial distance, r from the nucleus in any direction. Here, $\alpha_1 \alpha_2$ is the potential energy curve to the right of the nucleus. The dashed curve is to the left of nucleus. The potential energy at every point in space requires a '4D' picture. We require 3D for space coordinates and the 4th dimension for the potential energy axis. We herewith plot 'U' along some chosen axis (line) through the crystal, say, a row of ions.

To build up the potential energy distribution at any point, consider the first 'two' adjacent ions neglecting others. This construction is shown in Fig. 1.11 in which $\alpha_1 \alpha_2$ is the potential energy curve for α . Also, $\beta_1 \beta_2$ represents the potential energy curve for β . The resultant potential energy curve due to α and β is the sum of the two curves $\alpha_1 \alpha_2$ and $\beta_1 \beta_2$ i.e., the curve $\alpha_1 d\beta_1$ since ab + ac = ad.

The resultant curve is nearly the same as the original one but is lower and flatter than the individual curves. The flat region is the plateau region which is the dark curve in Fig. 1.11.







Fig. 1.11 Potential energy curve resulting from two nuclei α and β

1.28 Physics for Electronics Engineering and Information Science

Consider an entire row of nuclei α , β , r, δ , ε ,... etc., in the metallic lattice. The form of potential energy as we proceed along this line from one nucleus to the other until the metal surface is reached is shown in Fig. 1.12.

It can be easily seen that the potential varies appreciably in the immediate vicinity of the nuclei, but is approximately a constant for a very large volume of the metal indicated by the slowly varying portions i.e., the plateau region PQ in the region between the ions. The exact position of the metal surface cannot be defined. It is located at a small distance from the last nucleus ' ε ' in the row. There is no nucleus to the right of ' ε ' and hence, there is no lowering or flattering of the potential energy curve. Thus, a potential energy barrier exists at the surface of the metal.



barrier exists at the surface of the metal. Consider the motion of an electron at location A as shown in Fig. 1.12. This collides with and rebounds from the potential walls at a and b. Due to its close proximity to the nucleus, it cannot drift too much from it. Hence, an electron in such a location (inner shells) is strongly bound to the nucleus and does not take part

It is clear that for an electron in a level such as B (valence electrons), at no point within the metal, its energy is converted to potential energy. Hence, its velocity is not zero at any point and the electron travels almost freely throughout the metal. But when the electron reaches the surface of the metal, it encounters the potential energy barrier at C and rebounds back into the metal. An electron having energy as D (may have been obtained by the electron from outside source like light), does not collide with any potential walls, not even the one at the surface and hence is capable of leaving the metal.

Thus, the free electrons find themselves in a plateau or equipotential region and are subject to no forces (as

we take V = 0 in Schrödinger's equation for particle in a box). Omitting all the potential variations except the potential barrier at the surface, a simplified form of the potential energy picture of Fig. 1.12 is shown in Fig. 1.13. Here, the plateau region is taken as *zero*. This is valid because we are only interested in differences in potential.

in conduction.

The region outside the metal is now at a potential equal to E_B i.e., the height of the potential energy barrier is E_B (eV).

To escape from the metal, the electrons should overcome the barrier height E_B . It is seen from the particle in a box problem that even the free





electrons in the metal have their energies quantized. The allowed energy levels have been calculated. But how are the electrons distributed among these energy levels? i.e., how many electrons are available per unit volume in the range E to E + dE and hence, what is the total number of free electrons per unit volume. In other words, this means that we need to obtain the energy distribution of electrons. The relationship which would give the distribution is termed as the energy distribution function. The density of electrons i.e., the number of electrons per m^3 may be written as

$$dn_E = \rho_E dE \tag{1.72}$$

where dn_E is the number of free electrons per m^3 whose energies lie in the range E to E + dE and ρ_E is the density of electrons in this interval.

As there are no potential variations, we expect the same number of electrons in each cubic meter of the metal i.e., the density of electrons in space in electrons per m³ is a constant. But in each unit volume of the metal, there will be electrons with all possible energies.

i.e.,

 ρ_E = Number of electrons / m³ of metal

Hence,

$$= \int dn_E = \int \rho_E dE \tag{1.73}$$

where the integral sums over the entire energy range and n gives the total number of electrons per m³ of the metal. Equation (1.73) is represented in Fig. 1.14.

n =

Also, ρ_E may be written as

$$\rho_E = f(E)N(E) \tag{1.74}$$



where N(E) is known as the *density of states* i.e., number of states per eV per m³ in the conduction region and F(E)is the probability that a quantum state with emergy E is seen

Fig. 1.14 Illustration of Eqn. (1.73)

is the probability that a quantum state with energy E is occupied by an electron.

1.10.1 Expression for Density of Energy States

From the particle in a box problem, it is known that the number of energy states with a particular value of *E* depends on how many combination of quantum numbers result in the same value of '*n*'(Degeneracy), where $n = n_x^2 + n_y^2 + n_z^2$.

To calculate the number of energy states with all possible energies, construct a 3 dimensional sphere as shown in Fig. 1.15 (a). Here, sphere is chosen since it can represent degenerate states. Every point (n_1, n_2, n_3) within this space represents an energy state. As n_1 , n_2 and n_3 can have only positive values, only the positive

octant of the sphere should be considered. The volume of this octant is $\frac{1}{8} \times \frac{4}{3} \pi n^3$. As the ground state value

of (n_1, n_2, n_3) is (111), i.e., the ground state contains only one quantum state, it follows that a volume of 1^3 contains one state.



Fig. 1.15 Construction of sphere to derive the density of energy states (a) 3D structure and (b) 2D structure

Similarly a volume of 2^3 i.e., 8 cubic units contain 8 states as follows:

 $\{(111), (222), (112), (121), (211), (221), (122), (212)\}$

The quantum number (n_x, n_y, n_z) is a positive integer, it occupies only the positive octant of the sphere of radius 'n'. Hence, the available energy states within the volume of $\frac{1}{8} \times \frac{4}{3}\pi n^3$ i.e., $\frac{\pi}{6}n^3$ states.

To calculate the number of permissible states within a small interval E and E + dE, construct two spheres of radius n and (n + dn) as shown in Fig. 1.15(b). The number of states having energy values between E and E + dE is

$$N'(E) dE = \frac{\pi}{6} (n + dn)^3 - \frac{\pi}{6} n^3$$

Neglecting higher powers of *dn*, we get

$$N'(E)dE = \frac{\pi}{2}n^2 dn$$
 (1.75)

It is known from the particle in a box problem that the energy of an electron in a cubical box of length a and quantum state n is given by

$$E_n = \frac{n^2 h^2}{8 m a^2}$$
(1.76)

$$n^2 = \frac{8 \ ma^2 E}{h^2} \tag{1.77}$$

$$n = \sqrt{\frac{8 \ ma^2 E}{h^2}} \tag{1.78}$$

Differentiating Eqn. (1.77) with respect to n,

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

$$dn = \frac{1}{2n} \frac{8 ma^2}{h^2} dE$$

$$= \left[\frac{8 ma^2}{h^2}\right] \left[\frac{1}{2}\right] \sqrt{\frac{h^2}{8 ma^2}} \frac{dE}{\sqrt{E}}$$

$$= \frac{1}{2} \sqrt{\frac{8 ma^2}{h^2}} \frac{dE}{\sqrt{E}}$$
(1.79)

Substituting for Eqn. (1.77) and Eqn. (1.79) in Eqn. (1.75), we have

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

i.e.,

Electrical Properties of Materials 1.31

$$N'(E)dE = \frac{\pi}{4} \left[\frac{8 \ ma^2}{h^2} \right]^{3/2} \sqrt{E} \ dE \tag{1.80}$$

Now, by Pauli's exclusion principle, two electrons of opposite spin can occupy each state and hence, assuming that each state is for one electron, the number of states available for electron occupancy is given by

$$N'(E)dE = 2 \times \frac{\pi}{4} \left[\frac{8ma^2}{h^2} \right]^{3/2} \sqrt{E} \, dE$$
$$N'(E)dE = \frac{4\pi}{h^3} (2m)^{3/2} a^3 \sqrt{E} \, dE$$
(1.81)

i.e.,

The density of states i.e., number of energy states per unit volume in the range E to E + dE, is given by

$$N(E)dE = \frac{N'(E)dE}{a^3} = \frac{4\pi}{h^3} (2m)^{3/2} \sqrt{E} dE$$
(1.82)

where E is in joules.

Suppose *E* is in eV,then

$$N(E)dE = \frac{4\pi}{h^3} (2m)^{3/2} \sqrt{E} (1.6 \times 10^{-19})^{3/2} dE$$
(1.83)

or

$$N(E)dE = r\sqrt{E}dE \tag{1.84}$$

where
$$r = \frac{4\pi}{h^3} (2m)^{3/2} (1.602 \times 10^{-19})^{3/2}$$
.

1.11 FERMI-DIRAC STATISTICS

f(E) is the Fermi Dirac probability function which specifies the fraction of all states at energy *E* occupied under the conditions of thermal equilibrium. As electrons belong to a class of particles called 'Fermions', we can apply the Fermi Dirac statistics in this case. According to the above statistics,

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$$
(1.85)

where T is the temperature in Kelvin, E_F is the Fermi level energy or the characteristic energy for the crystal in eV or joules and k is the Boltzmann's constant in eV/K or joules/K. The value of Boltzmann's constant is $k = 1.38 \times 10^{-23}$ J/K.

The implications of the Fermi-level are as follows:

(i) For T = 0 K and for $E > E_F$, the Fermi-level is

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}} = \frac{1}{1 + e^{\left(\frac{E - E_F}{0}\right)}}$$

$$=\frac{1}{1+e^{\infty}}=\frac{1}{1+\infty}=\frac{1}{\infty}=0$$

Hence, for $E > E_F$ and at T = 0 K,

$$f(E) = 0 \tag{1.86}$$

This implies that all quantum states with energies greater than E_F are empty at T = 0 K.

(ii) For T = 0 K and for $E > E_F$, the Fermi-level is

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}} = \frac{1}{1 + e^{\left(\frac{-(E_F - E)}{kT}\right)}}$$
$$= \frac{1}{1 + e^{\left(\frac{(E_F - E)}{kT}\right)}} = \frac{1}{1 + \frac{1}{e^{\infty}}} = \frac{1}{1 + \frac{1}{\infty}} = \frac{1}{1 + 0} = 1$$

Hence, for $E > E_F$ and at T = 0 K,

$$f(E) = 1$$
 (1.87)

This implies that all quantum states with energies lesser than E_F will be occupied at T = 0 K. (iii) For any temperature and for $E = E_F$, the Fermi-level is

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}} = \frac{1}{1 + e^0} = \frac{1}{1 + 1} = \frac{1}{2}$$

Hence, for $E = E_F$ and at $T \neq 0$ K,

$$f(E) = \frac{1}{2}$$
(1.88)

So, the Fermi level can be referred to as the state with 50% probability of being filled at any temperature or it represents the level up to which electrons are filled at 0K. The above Fermi level equations can be represented graphically as shown in Fig. 1.16.



Fig. 1.16 f(E) vs E(a) for T = 0 K and (b) T > 0 K

A graph of ρ_E versus *E* is shown in Fig. 1.17.



From Fig. 1.17, it can be seen that even for T = 2500 K, only electrons near the level E_F acquire higher energies. The arrangement of electrons in a metal for T = 0 K and T > 0 K can be depicted as in Fig. 1.18.



Fig. 1.18 Arrangement of electrons in a metal for (a) T = OK, (b) T_1K and (c) $T_2K > T_1K$

From Fig 1.18, it is seen that only the electrons near the Fermi level will take part in the conduction process on application of an external field. This explains the lower value of conductivity obtained experimentally as against the higher value predicted by the classical free electron theory.

This is because of the wrong assumption that all the *n* free electrons take part in the conduction process and also possess same energy. But it is seen that in the quantum mechanical picture, even free electrons have an energy distribution and are quantized. It is now possible to calculate the total number of free electrons per m^3 in a metal at T = 0 K. The area under the curve of Fig.1.17 gives the total number of free electrons per m^3 i.e., 'n'.

Hence,

$$n = \int_{0}^{E_{F}} f(E)N(E)dE \quad \text{or} \qquad (\text{since only upto level } E_{F} \text{ is filled at } T = 0 \text{ K})$$
$$= \int_{0}^{E_{F}} 1 \times r\sqrt{E}dE = \frac{2}{3}rE_{F}^{3/2} \qquad (\text{since } f(E) = 1 \text{ for } E < E_{F})$$

Therefore,

$$n = \frac{2}{3} r E_F^{3/2} \tag{1.89}$$

where $r = \frac{4\pi}{h^3} (2m)^{3/2} (1.602 \times 10^{-19})^{3/2}$ with E_F in eV.

Hence, the Fermi level is

$$E_F = \left(\frac{3n}{2r}\right)^{2/3} \tag{1.90}$$

Making use of the value of r with E in eV as 6.82×10^{27} , we get

$$E_F = 3.64 \times 10^{-19} \, n^{2/3} \tag{1.91}$$

and

$$n = (4.48 \times 10^{27}) E_F^{3/2} \tag{1.92}$$

As *n* varies from metal to metal, E_F also varies.

In general, the number of carriers per unit volume is given by

$$n_c = \frac{4\pi}{h^3} (2m)^{3/2} \int_{\substack{\text{energy}\\\text{band}}} \frac{\sqrt{E} dE}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$$

where E is in joules.

Suppose E is in eV, then

$$n_c = r \int_{\substack{\text{energy}\\\text{band}}} \frac{\sqrt{EdE}}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$$
(1.93)

EXAMPLE 1.18

The specific gravity of tungsten is 18.8 and its atomic weight is 184. Assume that there are two free electrons per atom. Calculate the numerical value of n and E_F .

Solution Given specific gravity = 18.8, atomic weight = 184 and number of free electrons per atom = 2. The number of free electrons per m^3 is

$$n = 6.02 \times 10^{23} \frac{\text{molecules}}{\text{mole}} \times \frac{1 \text{ mole}}{184 \text{ g}} \times 18.8 \frac{\text{g}}{\text{cm}^3} \times \frac{2 \text{ electrons}}{\text{atom}} \times \frac{1 \text{ atom}}{\text{molecule}}$$
$$= \frac{6.02 \times 10^{23} \times 18.8 \times 2}{184 \times 10^{-6}} = 1.23 \times 10^{29} \text{ electrons/m}^3$$

Here, Avogadro's number equals 6.02×10^{23} molecules per mole.

The Fermi level is

$$E_F = 3.64 \times 10^{-19} n^{2/3} = 3.64 \times 10^{-19} \times (1.23 \times 10^{29})^{2/3} = 8.6 \text{ eV}$$

EXAMPLE 1.19

Using Fermi function, evaluate the temperature at which there is 1% probability of an electron in a solid having energy 0.5 eV above E_F of 5 eV.

Solution Given E = 5.5 eV, $f(E) = 0.01 \text{ and } E_F = 5 \text{ eV}$.

Electrical Properties of Materials 1.35

Also,
$$E - E_F = 0.5 \text{ eV} = 0.5 \times 1.6^{-19} = 0.8 \times 10^{-19} \text{ J}$$

1

The Fermi function, $f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}} = 0.01$

i.e.,

$$\frac{1}{1+e^{\left(\frac{E-E_F}{kT}\right)}} = \frac{1}{1+e^{\left(\frac{0.8\times10^{-19}}{1.38\times10^{-23}\times T}\right)}} = \frac{1}{100}$$
$$e^{\left(\frac{0.8\times10^{-19}}{1.38\times10^{-23}\times T}\right)} = 99$$

or

Taking natural log on both sides, we get

$$\frac{0.8 \times 10^{-19}}{1.38 \times 10^{-23} \times T} = \ln 99$$

1

Hence,

$$T = \frac{0.8 \times 10^{-19}}{1.38 \times 10^{-23} \times 4.595} = 1.261 \times 10^3 \text{ K}$$

EXAMPLE 1.20

Using Fermi Distribution function, obtain the value of f(E) for $E - E_F = 0.01$ eV at 200 K.

Solution Given $(E - E_F) = 0.01$ eV and T = 200 K.

The Fermi Distribution function is

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}} = \frac{1}{1 + e^{\left(\frac{0.01 \times 1.602 \times 10^{-19}}{1.38 \times 10^{-23} \times 200}\right)}} = \frac{1}{1 + e^{\left(\frac{0.01602 \times 10^{-19}}{276 \times 10^{-23}}\right)}}$$
$$= \frac{1}{1 + e^{0.58}} = \frac{1}{1 + 1.79} = 0.358 \text{ or } 35.8\%$$

EXAMPLE 1.21

If the energy level is lying 0.01 eV above Fermi level in a solid, what is the probability of this level being occupied by an electron at 270? (AU June 2010)

Solution Given $(E - E_F) = 0.01$ eV, T = 270 K and Boltzmann constant, $k = 1.38 \times 10^{-23}$ J/K.

The Fermi Distribution function is

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}} = \frac{1}{1 + e^{\left(\frac{0.01 \times 1.602 \times 10^{-19}}{1.38 \times 10^{-23} \times 270}\right)}}$$

$$=\frac{1}{1+e^{0.4294}}=\frac{1}{1+1.5363}=\frac{1}{2.5363}=0.3943$$
 or 39.43%

EXAMPLE 1.22

Calculate the Fermi energy and Fermi temperature in a metal. The Fermi velocity of electrons in the
metal is 0.86×10^6 m/s.(AU May 2011)

Solution Given Fermi velocity $v_F = 0.86 \times 10^6$ m/s.

(i) Fermi Energy,

$$E_F = \frac{1}{2}mv_F^2 = \frac{1}{2} \times 9.11 \times 10^{-31} \times (0.86 \times 10^6)^2$$
$$= \frac{6.736 \times 10^{-19}}{2} = 3.368 \times 10^{-19} \text{ J}$$

or

$$E_F = \frac{3.368 \times 10^{-19}}{1.602 \times 10^{-19}} = 2.105 \text{ eV}$$

(ii) Fermi temperature,

$$T_F = \frac{E_F}{k} = \frac{3.368 \times 10^{-19}}{1.38 \times 10^{-23}} = 24405 \text{ K}$$

EXAMPLE 1.23

Calculate the temperature at which there is 1% probability of a state with energy 0.5 eV above Fermi energy. (AU June 2011)

Solution Given $E - E_F = 0.5$ eV and f(E) = 1% = 0.01.

We know that

$$f(E) = \frac{1}{1 + e^{\frac{(E - E_F)}{kT}}}$$

$$1 + e^{\frac{(E - E_F)}{kT}} = \frac{1}{f(E)}$$
$$e^{\frac{(E - E_F)}{kT}} = \frac{1}{f(E)} - 1$$

i.e.,

Taking logarithm on both sides

$$\frac{E - E_F}{kT} = \log_e \left[\frac{1}{f(E)} - 1 \right]$$

Therefore,

$$T = \frac{(E - E_F)}{k \left[\log_e \left(\frac{1}{0.01} - 1 \right) \right]} = \frac{0.5 \times 1.602 \times 10^{-19}}{1.38 \times 10^{-23} \times \left[\log_e \left(\frac{1}{0.01} - 1 \right) \right]}$$
$$= \frac{8 \times 10^{-20}}{1.38 \times 10^{-23} \times \log_e 99} = \frac{8 \times 10^{-20}}{6.341 \times 10^{-23}} = 1261.6 \text{ K}$$

EXAMPLE 1.24

Calculate the number of states per unit volume in an energy interval of 0.01 eV above the Fermi energy of sodium metal. The Fermi energy of sodium at 0 K = 3 eV. (Given Planck's constant $h = 6.62 \times 10^{-34} \text{ J} \cdot \text{s}$). (AU June 2012)

Solution Given $E_F = 3 \text{ eV}$ and $h = 6.62 \times 10^{-34} \text{ J} \cdot \text{s}$

Also, Fermi energy of sodium at 0 K is

$$E_1 = E_F = 3 \text{ eV} = 3 \times 1.602 \times 10^{-19} = 4.8 \times 10^{-19} \text{ J}$$

and

$$E_2 = 3 \text{ eV} + 0.01 \text{ eV} = 3.01 \text{ eV} = 3.01 \times 1.602 \times 10^{-19} = 4.816 \times 10^{-19} \text{ J}$$

The number of energy states in unit volume having energies lying between E_1 and E_2 is

$$n = \frac{4\pi}{h^3} (2m)^{3/2} \int_{E_1}^{E_2} E^{1/2} dE$$

= $\frac{4\pi}{h^3} (2m)^{3/2} \times \frac{2}{3} \Big[E_2^{3/2} - E_1^{3/2} \Big]$
= $\frac{4 \times 3.14}{(6.62 \times 10^{-34})^3} (2 \times 9.1 \times 10^{-31})^{3/2} \times \frac{2}{3} \Big[(4.816 \times 10^{-19})^{3/2} - (4.8 \times 10^{-19})^{3/2} \Big]$
= $1.197 \times 10^{26} \text{ m}^{-3}$

Hence, the number of states per unit volume, $n = 1.197 \times 10^{26} \text{ m}^{-3}$

EXAMPLE 1.25

Assuming all the electrons to be free, calculate the total number of states below E = 5 eV in a volume 10^{-5} m³.

Solution Given $E_F = 5 \text{ eV}$.

The total number of free electrons per m³ is

$$n = \frac{2}{3} r E_F^{3/2}$$

where $r = \frac{4\pi}{h^3} (2m)^{3/2} (1.602 \times 10^{-19})^{3/2}$ with E_F in eV.

So,

$$r = \frac{4\pi}{h^3} (2m)^{3/2} (1.602 \times 10^{-19})^{3/2}$$

$$=\frac{4\pi}{(6.627\times10^{-34})^3}\times(2\times9.11\times10^{-31}\times1.602\times10^{-19})^{3/2}$$
$$=6.81\times10^{27}$$

Therefore,

$$n = \frac{2}{3}rE_F^{3/2} = \frac{2}{3} \times 6.81 \times 10^{27} \times (5)^{3/2}$$
$$= 5.07 \times 10^{28} \text{ electron/m}^3$$

Number of states in a volume of 10^{-5} m³ = 5.07 × 10^{28} × 10^{-5} = 5.07 × 10^{23}

EXAMPLE 1.26

The Fermi energy of silver is 5 eV, calculate the number of energy states for the free electrons in a cubical box of side 0.02 m lying below energy of 3 eV.

Solution Given the Fermi energy of silver, $E_F = 5 \text{ eV}$.

The total number of free electrons per m³ is

$$n = \frac{2}{3}rE_F^{3/2}$$

where $r = \frac{4\pi}{h^3} (2m)^{3/2} (1.602 \times 10^{-19})^{3/2}$ with E_F in eV.

So,

$$r = \frac{4\pi}{h^3} (2m)^{3/2} (1.602 \times 10^{-19})^{3/2}$$

$$=\frac{4\pi}{(6.627\times10^{-34})^3}\times(2\times9.11\times10^{-31}\times1.602\times10^{-19})^{3/2}$$
$$=6.81\times10^{27}$$

As we require states below 3 eV, we set $E_F = 3$ eV.

Therefore,

$$n = \frac{2}{3}rE_F^{3/2} = \frac{2}{3} \times 6.81 \times 10^{27} \times (3)^{3/2}$$

$$= 23.6 \times 10^{27}$$
 electrons/m³

Number of states in a volume of $0.02 \text{ m}^3 = 23.6 \times 10^{27} \times (0.02)^3 = 1.89 \times 10^{23}$

EXAMPLE 1.27

The atomic radius of silver which crystallizes in the FCC structure is 1.4 Å. Calculate E_F .

Solution Given atomic radius = 1.4Å = 1.4×10^{-10} m and the structure of silver is FCC. As the structure is FCC, there are 4 atoms/unit cell which implies 4 electrons per unit cell.

For FCC,
$$a = \frac{4 \times \text{radius}}{\sqrt{2}}$$

In $\left(\frac{4 \times \text{radius}}{\sqrt{2}}\right)^3$ m³, there are 4 electrons

In one m³, there are
$$\frac{4}{(4 \times 1.4 \times 10^{-10} / \sqrt{2})^3}$$
 electrons

Hence, number of electrons,
$$n = \frac{4}{\left(\frac{4 \times 1.4}{\sqrt{2}}\right)^3 \times 10^{-30}} = 6.44 \times 10^{28} / \text{m}^3$$

So, Fermi level is

$$E_F = \left(\frac{3n}{2r}\right)^{2/3}$$
 where $r = 6.81 \times 10^{27}$ (Refer to Example 1.25)

$$= \left(\frac{3 \times 6.44 \times 10^{28}}{2 \times 6.81 \times 10^{27}}\right)^{2/3}$$
$$= 0.365 \times 10^{-18} \times (64.4 \times 10^{27})^{2/3}$$
$$= 0.365 \times 16.07 = 5.87 \text{ eV}$$

EXAMPLE 1.28

Density of free electrons in aluminium is 18.1×10^{28} m⁻³. Calculate its Fermi energy at zero Kelvin. ($h = 6.62 \times 10^{-34}$ J·s and $m = 9.11 \times 10^{-31}$ kg) (AU Dec 2005)

Solution The Fermi energy of electrons at 0 K in joules is

$$E_F = \left(\frac{h^2}{8\,m}\right) \times \left[\frac{3N}{\pi}\right]^{2/3} = \frac{(6.62 \times 10^{-34})^2}{8 \times 9.11 \times 10^{-31}} \times \left[\frac{3 \times 18.1 \times 10^{28}}{3.14}\right]^{2/3}$$
$$= (6.0198 \times 10^{-38}) \times (1.7293 \times 10^{29})^{2/3} = 1.87 \times 10^{-18} \text{ J}$$

The Fermi energy of electrons at 0 K in eV is

$$E_F = \frac{1.87 \times 10^{-18}}{1.6 \times 10^{-19}} = 11.66 \text{ eV}$$

1.11.1 Mean Energy of the Electrons at 0 K

To find the mean energy of the electrons at 0 K, the total energy is calculated at 0 K i.e., $E_{\text{Total}(0)}$. Now the energy of electrons in the range *E* to *E* + *dE* is *E*. Total energy in the range *E* to *E* + *dE* is *f*(*E*)*N*(*E*)*dE*.*E*.

$$E_{\text{Total}(0)} = \int_{0}^{E_{F}} f(E)N(E)E.dE$$

~

Since f(E) = 1 at T = 0 K,

$$E_{\text{Total}(0)} = \int_{0}^{E_{F}} r\sqrt{E} \cdot E \cdot dE$$

= $\frac{2}{5} rE_{F}^{5/2}$ (1.94)

The mean energy is given by

$$\overline{E}_{0} = \frac{E_{\text{Total}(0)}}{n} = \frac{\frac{2}{5}rE_{F}^{5/2}}{\frac{2}{3}rE_{F}^{3/2}}$$

$$\overline{E}_{0} = \frac{3}{5}E_{F0}$$
(1.95)

For T > 0 K, it can be shown that,

$$\overline{E}_0 = E_{F0} \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{E_{F0}} \right) \right]$$
(1.96)

EXAMPLE 1.29

The Fermi energy of silver is 5.51 eV. What is the average energy of a free electron at 0 K? (AU May 2009)

Solution Given Fermi energy, $E_{F0} = 5.51 \text{ eV} = 5.51 \times 1.602 \times 10^{-19} \text{ J} = 8.827 \times 10^{-19} \text{ J}.$ The average energy is

$$\overline{E}_0 = \frac{3}{5} E_{F0} = \frac{3}{5} \times 8.827 \times 10^{-19} = 5.296 \times 10^{-19} \text{ J}$$
$$\overline{E}_0 = \frac{5.296 \times 10^{-19}}{1.602 \times 10^{-19}} = 3.306 \text{ eV}$$

or

Hence, average energy of a free electron at 0 K is 3.306 eV.

1.12 WORK FUNCTION

It is seen that the energies of free electrons in a metal are distributed from 0 to E_F . Hence, E_F is the maximum energy that is possessed by the electron. It is also seen that there exists a potential barrier near the metal surface. Both the above statements in the form of a graph are shown in Fig. 1.19.

Here, it is seen that the maximum energy of the electron at 0K is E_F . To escape, the minimum energy to be supplied is $E_w = E_B - E_F$ and is called the work function of the metal. Thus, work function can be defined as the minimum energy required to remove the electron from the metal surface at 0 K. Electrons with energy lesser than E_F will require a greater amount of energy to be removed. The maximum energy will be equal to the barrier height which would be the energy required to remove the electron in the lowest energy state. The excess energy thus absorbed by electrons in higher energy states appear as their kinetic energies in accordance with the equation $hf = hf_o + \frac{1}{2}mv^2$. This explains the concept of threshold frequency or energy and the energy distribution of electrons as that observed experimentally in Photoelectric effect.



Fig. 1.19 Energy distributions and existence of potential barrier

Successes of Quantum Free Electron Theory

- (i) This theory explains the specific heat capacity of materials.
- (ii) This theory yields resistivity $\rho \alpha T$ which is experimentally correct.
- (iii) This theory gives the correct mathematical expression for electrical and thermal conductivity of metals.
- (iv) Gives correct value of Lorentz number.

Failures of Quantum Free Electron Theory

- (i) This theory fails to distinguish between metal, semiconductor and insulator.
- (ii) It also fails to explain the positive value of Hall Co-efficient.
- (iii) Magnetic behaviour of some solids are not explained. For example, copper should be paramagnetic according to this theory but in reality copper is diamagnetic.

The above drawbacks result from the electron being assumed to move in a constant potential. But the true picture is that the electron moves in a periodic potential interacting with the crystal lattice which leads to the band theory of solids.

EXAMPLE 1.30

The energy required to remove an electron from sodium metal is 2.3 eV. Does sodium exhibit photoelectriceffect from an orange light having wavelength 2800Å?(AU Nov 2002, Nov 2003)

Solution Given $\lambda = 2800$ Å and the energy required to remove an electron from sodium i.e., $E_g = 2.3$ eV. Using orange light with wavelength of 2800 Å, the energy produced is

$$E = hf = \frac{hc}{\lambda}$$
 joules where $f = \frac{c}{\lambda}$

 $E = \frac{6.625 \times 10^{-34} \times 3 \times 10^8}{2800 \times 10^{-10} \times 1.602 \times 10^{-19}} = 4.4 \text{ eV}$

Therefore,

Since the energy produced (4.4 eV) is greater than the required value (2.3 eV), sodium exhibits photo-electric effect from an orange light having a wavelength of 2800 Å.

EXAMPLE 1.31

The photoelectric work function of aluminium is 4.2 eV. Calculate the kinetic energy of fastest photoelectron when light of wavelength 200 nm is incident on an aluminium surface. (1 eV = 1.6×10^{-19} J; $h = 6.625 \times 10^{-34}$ J·s and $c = 3 \times 10^8$ m/s) (AU Jan 2006)

Solution Given $\lambda = 200 \text{ nm} = 200 \times 10^{-19} \text{ V}$. and $\phi_0 = 4.2 \text{ eV} = 4.2 \times 1.6 \times 10^{-19} = 6.72 \times 10^{-19} \text{ V}$.

The kinetic energy is

$$KE_{\text{max}} = hf - \phi_0 = \frac{hc}{\lambda} - \phi_0$$
$$= \frac{6.625 \times 10^{-34} \times 3 \times 10^8}{200 \times 10^{-9}} - 6.72 \times 10^{-19} = 3.2 \times 10^{19} \text{ J}$$

The kinetic energy in terms of eV is

$$KE_{\text{max}} = \frac{3.2 \times 10^{-19}}{1.602 \times 10^{-19}} = 2 \text{ eV}$$

10

Therefore, the kinetic energy of fastest photoelectron is 2 eV.

EXAMPLE 1.32

The longest wavelength of incident light radiation that can eject photo electrons from platinum is 1980 Å. Calculate the work function of platinum in eV.

Solution Given $\lambda = 1980 \text{ Å} = 1980 \times 10^{-10} \text{ m}.$

The work function is

$$\phi = hf = \frac{hc}{\lambda} = \frac{6.625 \times 10^{-34} \times 3 \times 10^8}{1980 \times 10^{-10}} = 1.004 \times 10^{-18} \text{ J}$$

The work function in terms of eV is

$$\phi = \frac{1.004 \times 10^{-18}}{1.602 \times 10^{-19}} = 6.2736 \text{ eV}$$

EXAMPLE 1.33

If ultraviolet rays of 1000 Å are incident on molybdenum target, whose work function is 4.15 eV, find the maximum velocity of the ejected electrons.

Solution

Given $\lambda = 1000 \text{ Å} = 1000 \times 10^{-10} \text{ m}$ and the work function $\phi_0 = 4.15 \text{ eV}$.

The total energy of the photon is

$$hf = \phi_0 + \frac{1}{2}mv_{\max}^2$$

Therefore, $\frac{1}{2}mv_{\max}^2 = hf - \phi_0$

i.e.,

$$mv_{\max}^2 = 2\left(\frac{hc}{\lambda} - \phi_0\right)$$

Therefore, the maximum velocity of the ejected electrons is

$$v_{\text{max}} = \sqrt{\frac{2}{m} \left(\frac{hc}{\lambda} - \phi_0\right)}$$
$$= \sqrt{\frac{2}{9.11 \times 10^{-31}} \left(\frac{6.625 \times 10^{-34} \times 3 \times 10^8}{1000 \times 10^{-10}} - 4.15 \times 1.602 \times 10^{-19}\right)}$$
$$= \sqrt{(2.1954 \times 10^{30})(1.3235 \times 10^{-18})} = 1.704 \times 10^6 \text{ m/s}$$

EXAMPLE 1.34

The photo-electric work function for a surface is 2.4 eV. Light of wavelength 6800 Å shines on the given surface. Find the incident and threshold frequencies, also show whether there is a possibility for photoemission to take place.

Solution Given $\lambda = 6800 \text{ Å} = 6800 \times 10^{-10} \text{ m}$ and the work function $\phi_0 = 2.4 \text{ eV}$.

The work function, $\phi_0 = h f_0$

Hence, the threshold frequency, $f_0 = \frac{\phi_0}{h} = \frac{2.4 \times 1.602 \times 10^{-19}}{6.625 \times 10^{-34}} = 5.796 \times 10^{14} \text{ Hz}$

Incident frequency, $f = \frac{c}{\lambda} = \frac{3 \times 10^8}{6800 \times 10^{-10}} = 4.412 \times 10^{14} \text{ Hz}$

From the above calculated frequency values, it is seen that the incident frequency is less than the threshold frequency. So, there is no possibility for photoelectric emission to take place.

EXAMPLE 1.35

The work function of a surface is 1.15 eV. If a monochromatic source of wavelength 4000 Å is incident on it, calculate the maximum kinetic energy of the ejected electrons.

Solution Given $\lambda = 4000 \text{ Å} = 4000 \times 10^{-10} \text{ m}$ and the work function $\phi_0 = 1.51 \text{ eV}$.

The total energy of the photon is

$$hf = \phi_0 + \frac{1}{2}mv_{\max}^2$$

Therefore, the maximum kinetic energy of the ejected electrons is

$$\left(\frac{1}{2}mv_{\text{max}}^2\right) = hf - \phi_0 = \frac{hc}{\lambda} - \phi_0$$
$$= \frac{6.625 \times 10^{-34} \times 3 \times 10^8}{4000 \times 10^{-10}} - (1.51 \times 1.602 \times 10^{-19}) = 2.553 \times 10^{-19} \text{ J}$$

1.13 BAND THEORY OF SOLIDS

The X-ray diffraction studies reveal the ordered structure of a solid. In a solid, the atoms occupy the lattice sites and the spacing between the atoms is of the same order as that of the linear dimensions of atoms. Therefore, atoms in a solid interact strongly and set up an internal electric field, which is periodic in nature and affects the motion of free electrons. It is also seen that the wave nature of electrons and application of quantum mechanics to the motion of electrons in a solid lead to a sequence of closely spaced energy levels.

In 1928, Felix Bloch developed the band theory for the electrons moving in a periodic field provided by a crystal lattice which was later on simplified by Kronig and Penny in 1931. Knowledge of the formation of energy bands and the consequent restrictions imposed on electron motion in a solid are obtained from the band theory. These restrictions led to concepts like allowed and forbidden energy bands, effective mass of the electron and concept of holes. It must, however, always be remembered that in reality, no band exists as a physical entity. Bands are mathematical conclusions that result due to assumptions. Such considerations led to the invention of a gamut of solid state devices, which has revolutionized the field of electronics leading to miniaturization, micro-miniaturization and mass production of devices and systems.

1.14 ELECTRON IN A PERIODIC POTENTIAL

The behaviour of an electron in a periodic potential is described by constructing the electron wave functions using one-electron approximation. In this approximation, the total wave function for the system is obtained from the combination of wave functions, each one of which involves the coordinates of one electron. Thus, the net field experienced by an electron is assumed to be the resultant of the field due to the ion cores and the average field due to the charge distribution of all other electrons. This can be described by the Kronig-Penney model and the wave function associated with the electron is given by Bloch functions.

1.14.1 Bloch Theorem

In free electron theory, one assumes that an electron moves in constant potential V_0 leading to the Schrödinger equation for a one-dimensional case.

Electrical Properties of Materials 1.45

(1.102)

$$\frac{d^2\psi}{dx^2} + \left(\left(\frac{2m}{\hbar^2}\right)(E - V_o)\right)\psi = 0$$

i.e.,

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0 \tag{1.97}$$

The complete solution for Eqn. (1.97) is

$$\psi(x) = e^{\pm ik \cdot x} \cdot e^{\pm i\omega t} \qquad (\text{Type 1}) \tag{1.98}$$

where $\omega = \frac{E}{\hbar}$ and the kinetic energy, $E_{\text{kinetic}} = E - V_0 = \frac{\hbar^2 k^2}{2m}$

Solutions of the Type 1 represent the propagation of waves along *x*-axis.

Consider the Schrödinger equation for an electron moving in a one dimensional periodic potential. Also, let the potential energy of the electron satisfy the following equation.

$$V(x) = V(x+a) \tag{1.99}$$

where a is the period. The Schrödinger equation is then

$$\frac{d^2\psi}{dx^2} + \left(\left(\frac{2m}{\hbar^2}\right)(E - V(x))\right) = 0$$
(1.100)

There exists solution for this equation of the form,

$$\psi(x) = e^{\pm ik \cdot x} u_k(x) \qquad \text{(Type 2)} \tag{1.101}$$

where

$$u_k(x) = u_k (x+a).$$

i.e., the solutions are plane waves modulated by the function $u_k(x)$, which has the same periodicity as the lattice. This theorem is known as *Bloch theorem*. Functions of the Type 2 are called *Bloch functions*.

Now,

$$\psi(x) = e^{i\kappa \cdot x} u_k(x)$$
 has the property

$$\psi(x+a) = e^{\left[ik(x+a)\right]}u_k(x+a) = e^{ik\cdot a}\psi(x)$$

where $u_k(x + a) = u_k(x)$

Therefore, $\psi(x + a) = e^{\pm ik \cdot a} \psi(x)$

If the Eqn. (1.100) has solutions with the property (1.102), then the solution can be written as Bloch function and the theorem is proved.

1.14.2 Behaviour of an Electron in Periodic Potential: The Kronig–Penny Model

The essential features of the behaviour of electrons in a periodic potential may be illustrated with reference to a relatively simple model first discussed by Kronig and Penny. It is assumed that the potential energy of an electron has a form of a periodic array of square wells as indicated in the Fig. 1.20.



Fig. 1.20 One dimensional periodic potential

The period of the potential is a + b

The boundary conditions can be written as

$$V = 0 \quad \text{for } 0 < x < a$$

$$V = V_0 \quad \text{for } -b < x < 0 \tag{1.103}$$

The Schrödinger equations for the two regions are

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} E\psi = 0 \qquad \text{For } 0 < x < a \tag{1.104}$$

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)\psi = 0 \quad \text{For } -b < x < 0 \tag{1.105}$$

Here , we assume $E < V_0$

Also, defining, $\alpha^2 = \frac{2 mE}{\hbar^2}$ in Eqn. (1.104) and $\beta^2 = \frac{2 m}{\hbar^2} (V_0 - E)$ in Eqn. (1.105) and making use of $\psi(x) = e^{ikx} u_k(x)$, we have

$$\frac{d\Psi}{dx} = ike^{ikx}u_k(x) + ike^{ikx}u'_k(x)$$

$$\frac{d^2\Psi}{dx^2} = -k^2 e^{ikx}u_k(x) + ike^{ikx}u'_k(x) + ike^{ikx}u'_k(x) + e^{ikx}u''_k(x)$$
(1.106)

Substituting Eqn. (1.106) in Eqn. (1.104) and Eqn. (1.105), we get

$$\frac{d^2u}{dx^2} + 2ik\left(\frac{du}{dx}\right) + (\alpha^2 - k^2)u = 0 \qquad \text{for } 0 < x < a$$
$$\frac{d^2u}{dx^2} + 2ik\left(\frac{du}{dx}\right) - (\beta^2 - k^2)u = 0 \qquad \text{for } -b < x < 0$$

The solutions of these equations are

$$u_1 = Ae^{i(\alpha - k)x} + Be^{-i(\alpha + k)} \qquad 0 < x < a$$
$$u_2 = Ce^{i(\beta - k)x} + De^{-i(\beta + k)} \qquad -b < x < 0$$

where A, B, C and D are constants. These constants are chosen such that they satisfy the requirement of continuity of the wave function and their derivatives and the periodicity of $u_k(x)$ which leads us to the following condition.

i.e.,
$$\frac{\beta^2 - \alpha^2}{2\alpha\beta} \sinh\beta b \sin\alpha a + \cosh\beta b \cos\alpha a = \cos k(a+b)$$
(1.107)

Kronig and Penny make the assumptions that $V_0 \to \infty$ as $b \to 0$, but $V_0 b$ is finite and $\alpha^2 \ll \beta^2$. When $b \to 0$, sinh $\beta b = \beta b$ and cosh $\beta b = 1$.

Using the above assumptions, Eqn. (1.107) reduces to

$$\frac{\beta b \beta^2}{2\alpha\beta} \sin \alpha a + \cos \alpha a = \cos ka \tag{1.108}$$

Since $\beta^2 = \frac{2m}{\hbar^2} (V_0 - E)$, Eqn. (1.108) can be written as

$$\frac{mV_0b}{\hbar^2\alpha}\sin\alpha a + \cos\alpha a = \cos ka \tag{1.109}$$

Let $P = \frac{mV_0ab}{\hbar^2}$.

Therefore, Eqn. (1.109) will become

$$P\frac{\sin\alpha a}{\alpha a} + \cos\alpha a = \cos ka \tag{1.110}$$

for which solutions can be obtained.



Fig. 1.21 Graph of $\cos ka vs \alpha a$ for $P = 3\pi/2$

Figure 1.21 shows the graph of $\cos ka$ vs αa for $P = 3\pi/2$. The R.H.S of Eqn. (1.110) can accept only values between -1 and +1 as indicated by the horizontal line.

From Fig. 1.21, the following interesting conclusions may be drawn.

- (i) The energy spectrum of the electrons consists of a number of allowed energy bands separated by forbidden regions.
- (ii) The width of the allowed energy bands increases with increasing values of αa , i.e., with increasing energy.
- (iii) The width of a particular allowed band decreases with increasing *P* i.e., with increasing binding energy of the electrons. When $P \rightarrow \infty$, the allowed regions become infinitely narrow and the energy spectrum becomes a line spectrum. For, $P \rightarrow \infty$, Eqn. (1.110) has solutions only if $\sin \alpha a = 0$, i.e., $\alpha a = \pm n\pi$ with n = 1,2,3,..

According to this and equation (1.106), the energy spectrum is then given by

$$E_n = \frac{\pi^2 h^2}{2ma^2} \cdot n^2 \qquad \text{for } P \to \infty \tag{1.111}$$

which is the energy levels of a particle in a constant potential box.

The nearly free electron model and the Kronig–Penny model are compared by plotting a curve between the electron's energy and its momentum as shown in Fig. 1.22(a) and (b) respectively. From Fig. 1.22(a), a parabola is obtained which indicates that the energy varies continuously. But in Kronig Penny model, a parabola is obtained with some discontinuities as seen in Fig. 1.22(b). Thus, the analysis of the behaviour of an electron in the periodic lattice in the view of Kronig–Penney model and Bloch theorem yields the following results:

- (i) There are allowed energy bands separated by forbidden regions or band gaps.
- (ii) The electronic energy functions E(k) are periodic in the wave vector k.

The discontinuities in *E* shown in Fig. 1.22(b) occur at $ka = \pm n\pi$. i.e., at $k = \pm \frac{\pi}{a}, \pm \frac{2\pi}{a}, \pm \frac{3\pi}{a}, \dots$ The origin of the allowed energy bands and forbidden gaps is clear from Fig. 1.22(b).



Fig. 1.22 Energy vs momentum curve (a) Nearly free electron model and (b) Kronig–Penny model

1.15 CLASSIFICATION OF SOLIDS-METALS, SEMICONDUCTORS AND INSULATORS

Solids like conductors, semiconductors and insulators are classified based on the energy band gap. Band gap is the net gap considering all directions along which the electrons cannot take those values of energy that lie in that gap regardless of their direction of motion. But this band gap disappears when there is sufficient overlap in the energy bands for different directions.

In a solid, outermost band that is fully or partially filled is called the valence band. The band that is above the valence band and that is empty at 0K is called the conduction band. Solids can be classified on the basis of their band structure as conductors, semiconductors and insulators.



Fig. 1.23 Schematic band structure of conductors (a) Partially filled valence band and (b) Overlapping bands

Conductors are those solids which have vacant electron energy states immediately above the highest filled level of the valence band. This can happen in two ways. In the first case, the valance band in conductors is only partially filled as in Fig. 1.23(a). The electrons here can respond to an externally applied field by acquiring extra velocity and moving into higher energy states. In the second case, a full valence band overlaps the conduction band as shown in Fig. 1.23(b) so that the forbidden gap is zero. All metals are examples of conducting materials. Monovalent metals such as the alkali metals have one electron per atom in the outer most shell and the outer most energy bands are half filled in these metals. Divalent metals such as magnesium have overlapping conduction and valence bands. Therefore, they can also conduct even if the valence band is full. The band structure of trivalent metals such as aluminium is similar to that of monovalent metals.



Fig. 1.24 Band structures of semiconductor and insulator energy band of (a) Silicon and (b) Diamond

Semiconductors are those materials which have an energy gap of about 2 to 3 eV or less as in Fig. 1.24(a). Example: Silicon and Germanium. When the energy gap is 2eV or less, an appreciable number of electrons can be excited across the gap at room temperature. So, semiconductors conduct much better than insulators at room temperature but is still several orders of magnitude poorer than metals which have no forbidden gap. By adding impurities or by thermal excitation, the electrical conductivity in semiconductors can be increased.

1.50 Physics for Electronics Engineering and Information Science

Insulators are those materials which have an energy gap more than 3eV as in Fig. 1.24(b). Example: Glass, wood and diamond. It has been estimated that millions of volt/m of electrical potential would be necessary to accelerate an electron sufficiently to jump the forbidden gap. The other possibility for a transition is that electrons cross the gap by thermal excitation. At room temperature, the number of electrons that can be thermally excited across the gap in insulators such as diamond turns out to be extremely small.

So, the conductors are the materials having enormous electrical conduction; the insulators are the materials in which practically there is no electrical conduction and the semiconductors are the materials in which the electrical conduction is in between the electrical conduction of conductor and insulator.

1.16 ENERGY BANDS IN SOLIDS – TIGHT BINDING APPROXIMATION

As per the classical free electron theory, in a metal, there are fixed ion cores tightly bound to the lattice locations with the valence electrons of atoms free to move about the whole volume like the molecules of a perfect gas in a container and is called *free electron approximation*.

In the *tight binding approximation*, individual atoms are brought together one after another to build the solid. The electrons are *tightly bound* to the parent atoms initially. As they are brought closer to each other, allowed and forbidden energy bands are formed.

A solid may be imagined to be formed by allowing initially free atoms to gradually approach one another. As long as the atoms are widely separated, their interactions are negligible. Every atom has the same energy level diagram. The energy-level diagram for the entire system of N atoms resembles the energy-level diagram for a single atom; now each state of the system can be occupied by N electrons instead of just one. As the atoms come together to create a close packed periodic structure, they interact strongly due to their proximity to each other. By interaction, we mean that the positive nucleus of one atom attracts the electrons and repels the nucleus of the adjacent atom and vice versa. As a result, instead of one energy level which is the same for all N isolated atoms, there arise N closely spaced separate levels, which fall into groups. The energy levels are so closely spaced in the group that they form a virtual continuum, which is called an energy band.

Consider an imaginary situation where *N* hydrogen atoms approach each other to form solid hydrogen. Each hydrogen atom is characterized by one electron residing at 1s energy level corresponding to an atomic orbital ψ . As long as the separation *r* of two atoms *A* and *B* is much larger than the size *d* of the atoms ($r \gg d$) as seen in Fig. 1.25(a), the atoms do not interact.

1.16.1 Energy Level Splitting

When two atoms come close i.e., r = d, one energy splits into two energy levels as seen in Fig. 1.25(b). When three atoms approach each other closely, the original level splits into three levels; four atoms produce four levels and so on as shown in Fig. 1.25(c). In general, N interacting atoms cause a particular energy level to split into N levels. The group of energy levels resulting from splitting is so closely spaced that it is called an energy band as shown in Fig. 1.25(d).

The individual valence electrons no longer belong to individual atoms, but they now belong to all nuclei in the solid.



Fig. 1.25 Energy level splitting and band formation (a) Non-interacting atoms, (b) Two, (c) Three and (d) 'N' interacting atoms

A crystal (i.e., solid) consists of an enormous number of atoms arranged in a regular periodic structure. The extent of energy level splitting in the solid depends on the nearness of atoms in it. Let us assume that N identical atoms form the crystal. Figure 1.26 represents the energy level splitting in a solid as a function of interatomic distance. The energy levels of the isolated atoms are shown in Fig. 1.26(c). All the N atoms have identical sets of energy levels. The electrons fill the energy levels in each atom independently. Fig. 1.26(b) shows an atom sitting at the origin of the coordinate system.

Now let us imagine that other atoms approach this atom along the three axial directions and halt at the distance a_0 , which is the lattice constant of the crystal. As the atoms approach, a continuously increasing interaction occurs between the atoms. Each of the energy levels splits into many distinct levels and form energy bands, as shown in Fig. 1.26(b). Fig. 1.26(a) depicts the effect of slicing of Fig. 1.26 (b) at a_0 and it represents the energy band structure of the crystal. It is seen that corresponding to each allowed energy level of an isolated atom, there forms an allowed energy band; and that the allowed energy bands are separated by forbidden bands of energy.



Fig. 1.26 Energy level splitting in a solid as a function of interatomic distance (a) Energy band structure of the solid corresponding to the actual spacing of atoms in the solid, (b) Energy level splitting as a function of distance and (c) Discrete energy levels in an isolated atom

1.52 Physics for Electronics Engineering and Information Science

The degree of splitting of energy levels depends on their depth in the atom. The electrons in outer shells screen the core electrons belonging to inner shells. Consequently, the energy levels of inner shell electrons are split to a lesser degree. They form narrow core bands. They are always completely filled and do not participate in electrical conduction. In contrast, the energy levels of valence electrons are split more and form wider bands.

In general, N interacting atoms cause an energy level to split into (2l + 1)N levels. Thus, s-level (l = 0) splits into N levels whereas the p-level, consisting of three sublevels p_x , p_y and p_z , splits into 3N levels. Thus, in a solid each level of an individual atom splits into (2l + 1)N number of levels where N is the number of atoms in the system. Consequently, the maximum electron capacity of an s-band is 2N electrons whereas the capacity of a p-band is 6N electrons.

While occupying an energy band, electrons start from the lowest energy level in the band and the levels one after the other in the ascending order of energy. When 2N electrons occupy the N levels available in the band, the band is said to be *completely filled*. In case of non-availability of 2N electrons, the energy band gets *partially filled*. When there are no electrons to occupy the levels, the energy band remains *vacant*.

The width of an allowed or forbidden energy band is generally of the order of a few electron-volts. As *N* is very large, the energy separation between successive energy levels in an allowed band is very small and is of the order of 10^{-27} eV. At room temperature, the kinetic energy of the electrons of the order of $kT (\approx 0.027 \text{ eV})$ which is very large compared to the energy level separation in an allowed band. Consequently, electrons can easily move into higher vacant levels within the allowed energy band either due to thermal energy or due to a small externally applied electric field. On the other hand, electrons cannot jump across a forbidden band under normal thermal energy possessed by them or due to applied electric field. High temperatures are required to cause inter-band electron transitions.

1.16.2 Energy Band Diagram for Silicon

Silicon belongs to Group IV elements in the periodic table and has an atomic number 14. The electronic configuration of silicon atom is $1s^22s^22p^63s^23p^2$. It is seen that the inner *K* and *L* shells are closed and the corresponding bands would be completely filled. In the outer sub shells 3s and 3p, 3s -sub shell is closed. The 3p sub-shell is partially filled. Hence, it is expected to behave as a good conductor. But because of formation of a hybrid band, which later branches out, Silicon solid behaves as a semiconductor.

In the crystal formation process, when the atoms are very far apart, as at position 'd' in Fig. 1.27, the electrostatic interaction among them is negligible. Consequently, the electronic energy levels of the crystal will be the same as those of isolated atoms. As the separation between atoms decreases, the 3s and 3p split and two bands are formed, as shown at position 'c' in Fig. 1.27. The band corresponding to 3s level has N energy levels and the band corresponding to 3p level has 3N levels. Here, 2N electrons occupy N levels in 3p-band. It may now be noted that there is an energy gap between the two bands.

The energy gap is seen to decrease with the decrease in atomic spacing. At position 'b' in Fig. 1.27, the two bands merge and form a composite band. The 3N upper levels merge with N lower levels, giving rise to a total of 4N levels. These levels have to be occupied by the 4N electrons available in total, and so the lowermost 2N levels are filled. When the atomic distance in our imaginary crystal is further reduced, the interaction among the atoms becomes very strong. Beyond the lattice spacing 'b' in Fig. 1.27, it is found that the composite band branches out and once again two bands are formed, separated by a forbidden gap, E_{g} .



Fig. 1.27 Energy level splitting and band formation in crystals of Group IV A elements

The significant point is that the 4N energy levels are equally divided between the two branches. There is an equal distribution of levels, 2N in each of the two bands. The 4N electrons available in total at 3s and 3p levels, now occupy the lower energy band and leave the upper band vacant. The lower band constitutes the valence band and the upper band constitutes the conduction band. This is the situation at the actual spacing a_0' in the silicon crystal. This is normally referred as sp^3 hybridisation. At position a_0 , the two bands are not widely separated from each other. The value of E_g at 0K is 1.12 eV. At normal temperatures, a significant number of electrons will be thermally excited from valence band to conduction band. The electrons excited to conduction band respond to the external voltage and produce a modest flow of current. Thus, silicon behaves as a semiconductor.

It is also evident from Fig. 1.27 that the energy gap between the two branches goes on increasing with decreasing atomic distance. In case of diamond, the two bands are separated by 5.47 eV. Even at high temperatures, the thermal energy would be insufficient to excite enough number of electrons to the conduction band. Because of the non-availability of electrons in the conduction band, electrical conduction cannot take place in the material and hence, diamond behaves as an insulator.

This now brings a very important concept i.e., the effective mass of an electron and the concept of holes.

1.17 EFFECTIVE MASS OF ELECTRON

When an electron moves in a periodic potential, its rest mass can be replaced by the effective mass in all calculations. According to Newton's second law,

Force =
$$Mass \times Acceleration$$

i.e.,

F = ma

When the electron is treated as a wave within the semiconductor crystal, the convenient classical F = ma relation can be used to describe the motion of an electron under an applied force such as eE_x and the apparent mass of the electron in the crystal can be determined.

The velocity and acceleration of the electron in the conduction band are evaluated in response to an electric field E_x along -x direction that imposes an external force $F_{\text{ext}} = eE_x$ in the +x direction. Since the electron can be treated as a wave, the group velocity v_g , is obtained using $v_g = d\omega/dk$. The energy $E = \hbar\omega$ (ω is an "angular frequency" associated with the wave motion of the electron). Both E and ω depend on k.

Thus, the group velocity of the wave packet as shown in Fig. 1.28 is given by

$$v_{g} = \frac{1}{\hbar} \frac{dE}{dk}$$
(1.112)

$$\int_{K}^{\Psi \text{packet}} \int_{K}^{\Psi \text{packet}} \int_{K}^{\Psi \text{packet}} X$$

$$| \underbrace{fig. 1.28 \quad Wave \text{ packet}} |$$

In the presence of an electric field, the electron experiences a force $F_{ext} = eE_x$ from which it gains energy until, it collides with a lattice vibration. During a small time interval δt between collisions, the electron moves a distance $v_e \times \delta t$ and hence, gains energy δE , which is

$$\delta E = F_{\text{ext}} \times v_g \times \delta t \tag{1.113}$$

To find the acceleration of the electron and the effective mass, the above equation can be modified as $F_{\text{ext}} = m_e a$, where *a* is the acceleration.

Substituting Eqn. (1.112) in Eqn. (1.113), the relationship between the external force and energy is given by

$$F_{\text{ext}} = \frac{1}{v_g} \frac{dE}{dt} = \hbar \frac{dk}{dt}$$
(1.114)

Equation (1.114) is the reason for interpreting $\hbar k$ as the crystal momentum in as much as the rate of change of $\hbar k$ by the externally applied force.

The acceleration is given by

$$a = \frac{dv_g}{dt} \tag{1.115}$$

Electrical Properties of Materials 1.55

Substituting Eqn. (1.112) in Eqn. (1.115), we get

$$a = \frac{dv_g}{dt} = \frac{1}{\hbar} \frac{d^2 E}{dk} \frac{1}{dt} = \frac{1}{\hbar} \frac{d^2 E}{dk^2} \frac{dk}{dt}$$
(1.116)

Since $\frac{dk}{dt} = \frac{F_{\text{ext}}}{\hbar}$, the relationship between F_{ext} and a can be written as

$$F_{\text{ext}} = \frac{\hbar^2}{d^2 E / dk^2} a \tag{1.117}$$

Here, the response of a free electron to the external force is $F_{ext} = m_e a$, where m_e is the mass of free electron in vacuum. Therefore, it is quite clear from Eqn. (1.117), that the effective mass m_e^* of the electron in the crystal is

$$m_e^* = \frac{F_{\text{ext}}}{a} = \frac{\hbar^2}{d^2 E/dk^2}$$
(1.118)

Thus, the electron responds to an external force and moves as if its mass were given by Eqn. (1.118).

1.17.1 Effective Mass of Electrons in Metals

Generally, in most conductors, $m_e^* = m_e$, since the band is only partially filled. But the effective mass of electron in metals like Copper, Magnesium and Platinum is greater than the mass of free electron. So, if effective mass of electron is substituted instead of true free electron mass in the expressions for specific heat, electrical conductivity and thermal conductivity, the correct values can be obtained. The concept of effective mass is able to account for many experimental observations like high electronic specific heat of transition metals and their high paramagnetic susceptibilities.

1.17.2 Effective Mass of Electron in Semiconductors and Insulators

The effective mass plays an important role in the conduction process in semiconductors and insulators since they have full or almost filled valence bands. It is found that the effective mass m_e^* is negative near the zone edges of almost filled valence bands. Physically speaking, the electrons in these regions are accelerated in a direction opposite to the direction of the applied force. This is called the negative mass behaviour of electrons. The electrons with negative mass can be considered as a new entity having the same positive mass of that electron and the same positive charge as the numerical value of the electron's charge. The new entity is given the name *hole*.

1.18 CONCEPT OF HOLE

The advantage of the concept of positive holes is that the momentum and current of a nearly filled band with n empty states can be attributed to the presence of an equivalent number of n holes with the same positive mass and positive charge of that of electron. The holes are not real particles like electrons or positrons, but it is only a way of looking at the negative mass behaviour of electrons near the zone edge. Here, the motion of the effective negative mass electrons can be considered as the motion of the positive holes or positive vacant sites in a nearly full band and allow the electrons in the band to carry the current.

Figure 1.29 shows the positive hole conduction and effective negative electron mass conduction as equivalent situations. From Fig. 1.29, it is seen that the motion of electrons is opposite to that of the applied electric

field E whereas the movement of holes is in the direction of the field. Several phenomena like Hall effect, Thomson effect, etc. find ready explanation on the basis of the hole concept.



Fig. 1.29 Motion of electron in the conduction band and hole in the valence band with electric field 'E'

TWO MARK QUESTIONS AND ANSWERS

1. List out the three main theories developed for metals.

- (i) Classical free electron theory which is macroscopic and obeys all classical laws
- (ii) Quantum free electron theory which is microscopic and obeys all quantum laws
- (iii) Zone theory or Band theory which is microscopic and is based on energy bands of solids

2. What are the postulates of classical free electron theory?

- (i) Classical free electron theory visualizes a metal as an array of atoms or ions permeated by a gas of free electron.
- (ii) There is no mutual interaction between the free electrons or between ions and electrons.
- (iii) The free electrons can move randomly under the constant potential provided by the fixed ions of the lattice.
- (iv) When the field is applied, the free electrons move in the direction opposite to that of the field.
- (v) Due to the field applied, they acquire the velocity called *drift velocity*, and the electron velocity in the metal obeys Maxwell Boltzmann statistics.

3. What are the merits of classical free electron theory?

The merits of classical free electron theory are:

- (i) Verifies Ohm's law.
- (ii) Explains the electrical and thermal conductivity of metals.
- (iii) Used to derive Wiedemann-Franz law.
- (iv) The optical properties of metal can be explained using these metals.

4. What are the failures of classical free electron theory?

The failures of classical free electron theory are:

- (i) This is a macroscopic theory.
- (ii) Classical theory states that all free electrons will absorb energy but quantum theory states that only few free electrons will absorb energy.
- (iii) This theory cannot explain Compton, photoelectric effect, paramagnetism and ferromagnetism.

- (iv) The theoretical and experimental values of specific heat and paramagnetic susceptibility do not match with each other.
- (v) By classical theory, $K/\sigma T$ equal to constant for all temperature. But, by quantum theory, $K/\sigma T$ is not a constant for all temperatures.
- (vi) Lorentz number, by classical theory, does not have good agreement with the experimental values and is rectified by quantum theory.

5. Define electrical conductivity.

The electrical conductivity is defined as the quantity of electricity flowing per unit area per unit time maintained at unit potential gradient. It is denoted as σ and its unit is $\frac{1}{\Omega \cdot m}$ or $(\Omega \cdot m)^{-1}$.

6. Define mean free path.

The average distance traveled by free electrons between two successive collisions is called *mean free* path. It is represented by λ . i.e., $\tau_c = \lambda / \overline{v}$ where \overline{v} is the average velocity of random motion or r.m.s velocity of electrons.

7. What is drift velocity?

The drift velocity is defined as the average velocity acquired by free electrons in a particular direction due to the application of electric field and is denoted by v_d . i.e., $v_d = \mu E = (e\tau_c)/m$.

8. What is mobility of an electron?

The mobility is defined as the drift velocity gained by the electron per unit electric field strength and is represented by μ . i.e., $\mu = v_d/E$.

9. Define relaxation rime.

The time required for the electron to return back to its equilibrium state from its disturbed state after the application of electric field is known as *relaxation time*.

10. Define thermal conductivity.

The thermal conductivity is defined as the amount of heat conducted per unit area per unit time maintained at unit temperature gradient. It is denoted as *K* and its unit is $\frac{W}{m \cdot K}$ or $Wm^{-1}K^{-1}$.

11. What are the similarities between the electrical conductivity and thermal conductivity of metals?

- (i) The electrical conductivity and thermal conductivity decreases with increase in temperature and impurities.
- (ii) The electrical conductivity and thermal conductivity is high at low temperature.
- (iii) For non-metals, the electrical conductivity and thermal conductivity is very less.

1.58 Physics for Electronics Engineering and Information Science

S. No.	Drift velocity	Thermal velocity
1.	Drift velocity is the average velocity acquired by the free electrons in the presence of an electric field	Thermal velocity is the velocity acquired by an electron without the application of any electric field
2.	The electrons moving with drift velocity will be in a direction opposite to that of the field	The electrons moving with thermal velocity is random in nature.
3.	Drift velocity of electrons is very less	Thermal velocity of electrons is very high

12. What are the differences between drift velocity and thermal velocity of an electron?

13. Distinguish between electrical conductivity and thermal conductivity

S. No.	Electrical conductivity	Thermal conductivity
1.	The quantity of current flowing per unit area per unit time maintained at unit potential gradient. It is denoted as σ .	The amount of heat conducted per unit area per unit time maintained at unit temperature gradient. It is denoted as <i>K</i> .
2.	Purely due to number of free electrons	Due to both free electrons and phonons
3.	Conduction of electricity takes place from higher potential end to lower potential end	Conduction of heat takes place from hot end to cold end
4.	Unit: $\frac{1}{\Omega \cdot m}$ or $(\Omega \cdot m)^{-1}$	Unit: $\frac{W}{m \cdot K}$ or $Wm^{-1}K^{-1}$

14. Give the microscopic form of Ohm's Law in a metallic conductor. Whether the Ohm's Law is true at all temperatures?

The microscopic form of Ohm's Law in a metallic conductor is

 $J = \sigma E$

where J is the current density, σ is electrical conductivity and E is electric field intensity. The Ohm's Law is not true at all temperatures in a conductor, since the resistance of a conductor varies with temperature in a complicated manner at different range of temperature.

15. Define drift velocity. How is it different from thermal velocity of an electron?

The drift velocity is defined as the average velocity acquired by the free electron in a particular direction due to the application of the electric field *E*. It is given by $v_d = \mu E$.

The thermal velocity is random in nature and is very high (105 m/s). But the drift velocity is a directional one and is very small (50 cm/s)

16. State Wiedemann-Franz law.

The ratio of thermal conductivity to that of electrical conductivity is proportional to absolute temperature. It is given by

$$\frac{K}{\sigma} = LT$$

where the constant 'L' is called Lorentz number and is equal to $2.44 \times 10^{-8} \, W \cdot \Omega \cdot K^{-2}$.

17. What is Lorentz number?

The ratio of thermal conductivity (*K*) to that of electrical conductivity (σ) is proportional to absolute temperature.

i.e.,
$$\frac{K}{\sigma} = LT$$
 or $L = \frac{K}{\sigma T}$

The constant 'L' is called Lorentz number and is equal to $2.44 \times 10^{-8} \text{ W} \cdot \Omega \cdot \text{K}^{-2}$ at 293K.

18. What is Fermi-Dirac distribution function?

The probability f(E) of an electron occupying an energy level E is given by Fermi-Dirac distribution Function. It is given by

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$$

where E_F is the Fermi energy, k is Boltzmann constant and T is the temperature.

19. Define density of states. What is its use?

The density of states Z(E)dE is defined as the number of available electron states per unit volume in an energy interval (dE).

i.e.,

$$Z(E)dE = \left(\frac{\text{Number of States between } E \text{ and } E + dE}{\text{Volume of the metal piece}}\right)$$

It is used to determine Fermi energy at any temperature, mean energy, electron concentration etc.

20. What is Fermi level? Give its importance.

The maximum energy level to which electron can be filled is called Fermi energy. The probability of electron occupation is 50% at any temperature. It is the reference energy level, which separates the filled and vacant energy levels.

21. Define carrier concentration in a metal.

The carrier concentration in a metal is defined as number of filled energy states between the energy levels dE.

22. Define bound electron and free electron.

Bound electron: Inner shell electrons of an isolated atom are strongly bound to the nucleus through electrostatic attraction. They do not contribute to the conductivity of the metal.

Free electron: Detached valence electron, that can freely move in a random manner inside the metal. They contribute to the conductivity of the metal.

23. What are the sources of electrical resistance in metals?

Lattice defects and thermal vibrations of the lattice:

When the electron is moving through a perfect periodic lattice, there is no resistivity except temperature dependent resistivity. The impurities and residual defects produce so many scattering centers and

reduce the mean free path of electrons. Similarly, if the vibration amplitude increases with the increase of temperature, the mean free path of electron decreases. Hence, the resistivity increases.

24. What is the importance of Quantum Free Electron Theory?

- (i) In quantum free electron theory, the wave aspects of electrons is taken into account.
- (ii) Fermi level electrons are responsible for electrical conductivity and thermal conductivity. Hence, the correct values of electrical conductivity, thermal conductivity and electronic specific heat are obtained.

25. What are the merits of quantum free electron theory?

- (i) Quantum free electron theory treats the electron quantum mechanically rather than classically.
- (ii) This theory also explains electrical conductivity and thermal conductivity and specific heat capacity of metals, photoelectric effect and Compton effect etc.

26. What are the drawbacks of quantum free electron theory?

- (i) Even though Quantum free electron theory explains most of the physical properties of metals, it fails to provide the distinction between metals, semiconductors and insulators.
- (ii) It also fails to explain the positive value of Hall coefficient and some transport properties of metals.

27. Show that at 0K, E_F is the maximum filled energy level.

The energy level is given by

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$$

where E_F is the Fermi energy, T is the temperature and k is Boltzmann constant.

- (i) For T = 0 K and for $E < E_F$, the Fermi-level f(E) = 1. This implies that all quantum states with energies lesser than E_F will be occupied at T = 0 K.
- (ii) For T = 0 K and for $E > E_F$, the Fermi-level f(E) = 0. This implies that all quantum states with energies greater than E_F are empty at T = 0 K.

Therefore, at 0 K, E_F is the maximum filled energy level.

28. What is stopping potential?

The retarding potential difference required to just stop the most energetic photoelectron, which is emitted by the metal, is called stopping potential.

29. What is work function?

The amount of kinetic energy required at absolute zero temperature to move the electron from outer orbit is called work function.

30. Distinguish between Energy state, Energy level and Energy band.

Energy state: The energy state is described by wave function and four quantum numbers.

Energy level: The energy of the electron in a shell or orbit of an atom is expressed in terms of energy level. An energy level may contain one or more energy states.

Energy band: The band is formed only in crystals or solids. The closed packing of array of energy levels forms a band.

31. What are the classifications of solids based on Band Theory?

According to the width of the gap between the conduction and valence bands, solids can be classified broadly into three groups as follows:

- (i) Metals
- (ii) Semiconductors
- (iii) Insulators

32. What is an energy band?

The splitting up of energy levels, which are so closely placed, form a band of energy levels and it is known as energy band.

33. How will you classify the metal, semiconductor and insulator based on band theory?

Metal: No energy gap between conduction band and valence band.

Semiconductor: The band gap between completely filled conduction band and valence band are relatively narrow bands. The energy gap for semiconductor is $E_g = 1 \text{eV}$.

Insulators: The band gap between completely filled conduction band and valence band are relatively wider. The energy gap for insulator is $E_g = 3$ eV.

34. What is meant by electron theory of solids?

The electrons in the outer most orbits of the atoms, which constitute a solid, determine its electrical properties. The electron theory of solids explains the structure and properties of solids through their electronic structure.

REVIEW QUESTIONS

- 1. Mention few important postulates of free electron theory of metals.
- 2. Obtain an expression for thermal conductivity of metals based on classical free electron theory.
- 3. What are the drawbacks of classical free electron theory of metals?
- 4. Derive expression for electrical and thermal conductivities based on electron theory.
- 5. The density of silver is 10.5×10^3 kg/m³. The atomic weight of silver is 107.9. Assuming that each silver atom provides one conduction electron (i) calculate the density of free electrons. The conductivity of silver at 20° C is $6.8 \times 10^7 \Omega^{-1} m^{-1}$. (ii) Calculate the electron mobility. (Answer: 5.86×10^{28} ; $7.25 \times 10^{-3} m^2$ /V.s)
- 6. Explain the terms: (i) drift velocity and (ii) carrier mobility.
- 7. A rectangular block of solid is connected to a dc voltage source. Obtain the expression:
 - (a) for the current density flowing through the block and
 - (b) for the conductivity of the material in terms of the concentration of carrier in it.
- 8. State and prove Wiedemann-Franz law. List the drawbacks of classical free electron theory.
- 9. On the basis of free electron theory, derive an expression for electrical and thermal conductivity. Hence, deduce Wiedemann-Franz law.
- 10. Deduce expression for electrical conductivity of a conducting material and hence, obtain Wiedemann-Franz law.

1.62 Physics for Electronics Engineering and Information Science

- 11. The thermal and electrical conductivity of copper at 20°C are 380 Wm⁻¹K⁻¹ and 5.87 × 10⁷ ($\Omega \cdot m$)⁻¹ respectively.Calculate the Lorentz number.(Answer: $L = 2.2094 \times 10^{-8} \text{ W} \cdot \Omega \cdot \text{K}^{-2}$)
- 12. Copper has electrical conductivity at 300 K as 6.40×10^7 mho m⁻¹. Calculate the thermal conductivity of copper. (Answer: K = 468.48 Wm⁻¹K⁻¹)
- 13. Explain quantum free electron theory.
- 14. Derive an expression for density of states and calculate the carrier concentration in metals.
- 15. Write down an expression for the probability of occupancy of particular energy states of an electron in an intrinsic semiconductor. Represent it graphically at oK and room temperature.
- 16. Write Fermi-Dirac distribution function. Explain how Fermi function varies with temperature. Give its significance.
- 17. Derive an expression for Fermi level in conductors. Obtain the expression for carrier concentration in terms of Fermi level.
- 18. Define Fermi distribution function. Show that at all temperature (*T* > oK), probability occupancy of Fermi level is 50%.
- 19. What is Fermi level and Fermi energy?
- 20. At what temperature, can we except a 10% probability that electrons in silver have an energy which is 1% above the Fermi energy? The Fermi energy of silver is 5.5 eV.
- 21. Why is that only the electrons near the Fermi level contribute to electrical conductivity?
- 22. Write down the Fermi-Dirac equation for the probability of occupation of an energy level E by an electron. Show that the probability of its occupancy by an electron is zero if $E > E_F$ and unity if $E < E_F$ at temperature oK.
- 23. Write down Fermi distribution function f(E). Show graphically and analytically that f(E) as function of E always passes through a point ($E_F = 1/2$) at different temperature.
- 24. In a solid, consider the energy lying 0.01 eV above Fermi level. What is the probability of this level being occupied by an electron at 300 K? (Answer: 0.595)
- 25. In a solid, consider the energy lying 0.01 eV above Fermi level. What is the probability of this level being not occupied by an electron at 300 K? (Answer: 0.405)
- 26. Calculate Fermi energy and Fermi temperature in a metal. The fermi velocity of electron in the metal is 8.6×10^5 m/s. (Answer: 2.1 eV; 24.3 × 10³ K)
- 27. Calculate the probabilities for an electronic state to be occupied at 20°C if the energy of these states lies 0.11 eV above and 0.11 eV below the Fermi level. (Answer: 0.0126; 0.9874)
- 28. Why we have energy bands in solids but energy levels in gases?
- 29. According to band theory, a completely filled band or empty band is not associated with electrical conduction. Only partially filled is responsible for electrical conduction. Explain why?
- 30. Explain the concept of negative mass on the basis band theory.
- 31. What is meant by effective mass of electron?
- 32. Explain formation of energy bands in solids on the basis of band theory of solids.
- 33. Explain the classification of metals, semiconductors and insulators based on the band theory.
- 34. How does the band theory differ from the free electron model in explaining the properties of metals?
- 35. Explain energy band diagram of silicon showing a graph of variation of potential energy with distance. Explain semi conducting nature of the silicon. With similar band structure, why is diamond called insulator?
- 36. Draw a graph showing variation of electron energy in germanium crystals as a function of inter-atomic distance. Explain from it why germanium is an insulator at oK and semiconductor at 72°C.
- 37. Explain the concept of hole.
2

Semiconductor Physics

2.1 INTRODUCTION

In order to understand the voltage-current characteristics of a semiconductor, knowledge of electron behavior in a semiconductor is necessary when the electron is subjected to various potential functions. This chapter deals with the characteristics and concentration of electrons bound to an atom in semiconductors or bound within a finite space to be that the electron can take on only discrete values of energy. The carrier transport phenomena forms the basis for obtaining the voltage-current characteristics of semiconductor devices. The two basic carrier transport mechanisms in semiconductor materials i.e., drift and diffusion transport are presented in this chapter.

In any semiconductor device, there will be atleast one junction between *P*-type and *N*-type semiconductor regions. Since the characteristics and operation of semiconductor devices are connected to these *PN* junctions, considerable attention is given to the basic device i.e., *PN* junction diode in this chapter. This chapter also discusses the operation and characteristics of tunnel diode in detail. The contact of semiconductor devices or integrated circuits, with the outside world, is done through metal-semiconductor junctions or ohmic contacts. This chapter explains ohmic contact, a low-resistance junction, which provides the current flow in both directions. Schottky diode, Hall Effect and devices, MOS capacitor and power transistors are also discussed in this chapter.

2.2 ATOMIC ENERGY LEVEL DIAGRAM

In the energy level diagram, the discrete energy states are represented by horizontal lines, and the height of the line represents the total energy E_n . Figure 2.1 shows the energy level diagram for hydrogen. The number immediately to the right of a line gives the value of integer n, while the number to the left of each line gives the energy to this level in electron volts. The lowest energy level E_1 is called the normal or the ground state of the atom and the higher energy levels E_2 , E_3 , E_4 , ... are called the excited states. As n increases, the energy levels crowd and tend to form a continuum.



Fig. 2.1 Energy level diagram of hydrogen

Sometimes, it is more convenient to specify the emitted radiation by its wavelength, λ in Angstroms which can be written as,

$$\lambda = \frac{12,400}{E_2 - E_1} \tag{2.1}$$

where E_2 and E_1 are the energy levels in electron volts.

2.3 PRACTICAL SEMICONDUCTOR MATERIALS

All semiconductors have crystalline structure. The most commonly used semiconductor materials, germanium, silicon and gallium arsenide have practical applications in electronics. The most frequently used semiconductors are germanium and silicon because the energy required to break their covalent bonds and release a free electron from their valence bands is lesser than that required for gallium arsenide. The energy required for releasing an electron from the valence band is 0.66 eV for germanium, 1.08 eV for silicon and 1.58 eV for gallium arsenide.

Germanium can be purified relatively well and crystallised easily. Germanium is an earth element and it is obtained from the ash of the certain coals or from the flue dust of the zinc smelters. The recovered germanium is in the form of germanium dioxide powder which is then reduced to pure germanium. Germanium diodes are used as infrared detectors in fibre-optic communication system because of narrower energy gap.

Silicon is an element found in most of the common rocks. Sand is silicon dioxide which is then reduced to 100% pure silicon. Silicon dioxide is a natural insulator which is useful in the fabrication of semiconductor devices and integrated circuits. Silicon is largely preferred to germanium because of its large gap energy, which produces improved device properties at high temperatures. Silicon is a better thermal conductor and is required to remove unavoidable heat developed in the device.

Gallium arsenide has higher electron mobility, μ_n which leads to faster switching capabilities. It has high temperature operating capabilities because of its larger energy gap.

2.4 CLASSIFICATION OF SEMICONDUCTORS

Semiconductors are classified into many ways. Their classification is based on the (i) number of elements, (ii) purity, (iii) doping concentration and (iv) bandgap.

Based on the number of elements, the semiconductors are classified into elemental and compound type. Semiconductor made use of only one element is called *elemental semiconductor*. Example: Si and Ge. Semiconductor made use of more than one element is called *compound semiconductor*. Some examples of compound semiconductors are: two element (binary) compound semiconductors: GaAs, GaP, InP, InN (III and V group compounds), ZnS, ZnSi, CdS, CdSi (II and IV group compounds) and SiC, SiGe (IV and IV group compounds), three element (tertiary) compound semiconductors: GaAsP, AlGaAs, four element (quaternary) compound semiconductors: GaAlAsP.

Based on the purity of atoms, the semiconductors are classified into (i) intrinsic (pure) and (ii) extrinsic (impure) types. The extrinsic semiconductors are of *N*-type and *P*-type.

Intrinsic semiconductor A pure semiconductor is called intrinsic semiconductor. Even at room temperature, some of the valence electrons may acquire sufficient energy to enter the conduction band to form free electrons. Under the influence of electric field, these electrons constitute electric current. A missing electron in the valence band leaves a vacant space there, which is known as a *hole*. Holes also contribute to electric current. In an intrinsic semiconductor, even at room temperature, electron-hole pairs are created. When electric field is applied across an intrinsic semiconductor, the current conduction takes place due to free electrons and holes. Under the influence of electric field, total current through the semiconductor is the sum of currents due to free electrons and holes.

Extrinsic semiconductor Due to the poor conduction at room temperature, the intrinsic semiconductor as such is not useful in the electronic devices. Hence, the current conduction capability of the intrinsic semiconductor should be increased. This can be achieved by adding a small amount of impurity to the intrinsic semiconductor, so that it becomes impure or extrinsic semiconductor. This process of adding impurity is known as *doping*. The amount of impurity added is extremely small; say 1 to 2 atoms of impurity for 106 intrinsic atoms.

N-type A small amount of pentavalent impurities such as arsenic, antimony or phosphorus is added to the pure semiconductor (germanium or silicon crystal) to get *N*-type semiconductor. The addition of pentavalent impurity (arsenic, antimony or phosphorus) increases the number of electrons in the conduction band, thereby increasing the conductivity of *N*-type semiconductor. As a result of doping, the number of free electrons far exceeds the number of holes in an *N*-type semiconductor. Hence, electrons are called majority carriers and holes are called minority carriers.

P-type A small amount of trivalent impurity such as aluminium or boron is added to the pure semiconductor (germanium or silicon crystal) to get the *P*-type semiconductor. The addition of trivalent impurity (boron, aluminium or gallium) increases the number of holes in the valence band, thereby increasing the conductivity of *P*-type semiconductor. As a result of doping, the number of free holes far exceeds the number of electrons in a *P*-type semiconductor. Hence, holes are called majority carriers and electrons are called minority carriers.

Based on the doping concentration, the semiconductors are classified into (i) *Lightly doped* with doping concentration of 10^{13} to 10^{14} impurity atoms/cm³ (ii) *Moderately doped* with doping concentration of 10^{15} to 10^{16} impurity atoms/cm³ (iii) *Heavily doped* with doping concentration of 10^{17} to 10^{18} impurity atoms/cm³ and (iv) *Degenerative* with doping concentration greater than 10^{18} impurity atoms/ cm³.

Based on the bandgap, the semiconductors are classified into (i) direct bandgap and (ii) indirect bandgap type. If the momentum of electrons and holes is the same in both the conduction band and the valence band, then the semiconductor is called *direct bandgap* semiconductor. Here, an electron can directly emit a photon. Example: Gallium Arsenide (GaAs) and Indium Arsenide (InAs). In an *indirect bandgap* semiconductor, a photon cannot be emitted because the electron must pass through an intermediate state and transfer momentum to the crystal lattice. Example: Silicon (Si) and Germanium (Ge).

2.5 INTRINSIC SEMICONDUCTORS WITH ENERGY BAND DIAGRAM

A semiconductor in which holes and electrons are created solely by thermal excitation across the energy gap is known as *intrinsic semiconductor*. Pure crystals of Germanium (Ge) and Silicon (Si) are the two important elemental semiconductors used in electronic devices. Ge has a total of 32 electrons and Si has a total of 14 electrons in their atomic structures. Since there are four valence electrons in each of them, they are called *tetravalent atoms*. Both these elements are from Group IV in the periodic table. The neighbouring atoms form covalent bonds by sharing four electrons with each other so as to achieve inert gas or stable structure. In the case of pure (intrinsic) Ge, the covalent bonds have to be broken to provide conduction electrons. The energy required to break such a covalent bond is about 0.66 eV for Ge and 1.1 eV for Si at 300 K.



Fig. 2.2 Covalent Bond and Energy Band diagram for intrinsic semiconductor at (a)T = 0 K and (b)T > 0 K

Figure 2.2 shows the covalent bond and energy band diagram for intrinsic semiconductor. The covalent bond linkage of Ge can be seen in Fig. 2.2 (a) at 0 K. Here, the valence band is completely filled and the conduction band is completely empty. Thus, the intrinsic semiconductor behaves as a perfect electrical insulator.

The germanium crystal with a broken covalent band and the formation of electron-hole pair can be seen in Fig. 2.2 (b) for T > 0 K. When temperature increases above 0 K, the valence band electrons are thermally excited and shifted to the conduction band across the band gap. Hence, there are some free electrons in the conduction band and an equal number of holes or vacant sites in the valence band. As the temperature increases further, the generation of electron hole pairs is also increased and hence, the concentration of electrons in the conduction band and the concentration of holes are also increased. Therefore, at any temperature, the concentration of free electrons is always equal to the concentration of holes in an intrinsic semiconductor.

In equilibrium, the rate of generation of electron-hole pairs will be equal to their rate of recombination. Depending on the manner in which the recombination process takes place, the semiconductors can be classified as direct band and indirect bandgap semiconductors.

2.6 DIRECT AND INDIRECT BANDGAP SEMICONDUCTORS

The process of recombination differs in different semiconductors. This depends on the value of momentum, the position of the top of the valence band and the bottom of the conduction band occurs in the semiconductor. The minimum difference in energy between the two is termed as the band gap. In general, the minimum (bottom) of the conduction band and the maximum (top) of the valence band do not occur at the same value of momentum. If they occur at the same value of momentum, the semiconductor is called a *direct bandgap semiconductor* and if they occur at different values of momentum, it is termed as *indirect bandgap semiconductor*. The energy- momentum curves for direct and indirect bandgap semiconductors are shown in Figs. 2.3(a) and 2.3(b) respectively.



Fig. 2.3 Energy versus momentum curve for (a) direct band gap semiconductor and (b) indirect band gap semiconductor

In a direct band gap semiconductor, a photon of energy exactly equal to the band gap E_G can produce an electron-hole pair as the electron in the valence band need not be given additional momentum to move to the conduction band. But in an indirect band gap semiconductor, additional momentum must be provided to an electron in the valence band (as k_1 should be increased to k_2 first). Only then, a photon of energy E_G will be able to produce an electron-hole pair.

For acquiring this additional momentum, the electron apart from interacting with the photon to gain energy, should also interact with a lattice vibration called phonon. As this requires three entities, namely, an electron, a photon and a phonon to intersect, the indirect process takes a longer time. Thus, the recombination process

is more effective in a direct band gap semiconductor than in an indirect band gap semiconductor where the process requires mediation by a phonon.

Due to the above mentioned properties, gallium arsenide (GaAs) and other direct band gap semiconductors are used to make optical devices such as LEDs and semiconductor lasers, whereas silicon (Si), which is an indirect band gap semiconductor is used as a substrate for semiconductor device. Table 2.1 lists some of the semiconducting elements and compounds with their band gaps at T = 300K.

Semiconducting elements/Compounds	Material	Direct/Indirect Bandgap	Bandgap Energy at 300 K (eV)
Elements	Ge	Indirect	0.66
	Si	Indirect	1.12
	C (Diamond)	Indirect	5.47
	Sn (Grey Tin)	Direct	0.08
Group III-V compounds	GaAs	Direct	1.42
	GaP	Indirect	2.26
	GaN	Direct	3.36
	InAs	Direct	0.36
	InSb	Direct	0.17
	InN	Direct	0.70
Group II-VI compounds	ZnS	Direct	3.68
	ZnO	Direct	3.35
	CdSe	Direct	1.70
Group IV-IV compounds	α-SiC	Indirect	2.99

 Table 2.1
 List of some semiconducting elements and compounds together with their bandgaps at T = 300 K

2.7 CONDUCTIVITY OF SEMICONDUCTOR

In a pure semiconductor, the number of holes is equal to the number of electrons. Thermal agitation continues to produce new electron-hole pairs and the electron-hole pair disappears because of recombination. With each electron-hole pair created, two charge-carrying particles are formed. One is negative which is the free electron with mobility μ_n . The other is positive, i.e., the hole with mobility μ_p . The electrons and holes move in opposite directions in an electric field *E*, but since they are of opposite sign, the current due to each is in the same direction. Hence the total current density *J* within the intrinsic semiconductor is given by

$$J = J_n + J_p$$

= $q \cdot n \cdot \mu_n E + q \cdot p \cdot \mu_p \cdot E$
= $(n\mu_n + p\mu_p)qE$
= σE (2.2)

where J_n = electron drift current density

 J_p = hole drift current density

- n = number of electrons per unit volume, i.e., magnitude of free-electron (negative) concentration
- p = number of holes per unit volume, i.e., magnitude of hole (positive) concentration

- E = applied electric field strength, V/m
- q = charge of electron or hole, Coulomb

Hence, σ is the conductivity of a semiconductor which is equal to $(n \mu_n + p \mu_p)q$. The resistivity (ρ) of a

semiconductor is the reciprocal of conductivity, i.e., $\rho = \frac{1}{\sigma}$.

It is evident from the above equation that current density within a semiconductor is directly proportional to the applied electric field.

For pure (intrinsic) semiconductor, $n = p = n_i$, where n_i is the intrinsic carrier concentration.

Therefore, $J = n_i (\mu_n + \mu_p) qE$

and conductivity of an intrinsic semiconductor is $\sigma_i = q \cdot n_i(\mu_n + \mu_p)$. Hence, it is clear that conductivity of an intrinsic semiconductor depends upon its intrinsic concentration (n_i) and the mobility of electrons (μ_n) and holes (μ_p) . The intrinsic conductivity of germanium and silicon increase by approximately 5 per cent per °C and 7 per cent per °C rise in temperature respectively due to the influence of n_i .

EXAMPLE 2.1

The mobility of free electrons and holes in pure germanium are 3800 and 1800 cm²/V-s respectively. The corresponding values for pure silicon are 1300 and 500 cm²/V-s, respectively. Determine the values of intrinsic conductivity for both germanium and silicon. Assume $n_i = 2.5 \times 10^{13}$ cm⁻³ for germanium and $n_i = 1.5 \times 10^{10}$ cm⁻³ for silicon at room temperature.

Solution (i) The intrinsic conductivity for germanium,

$$\sigma_i = qn_i(\mu_n + \mu_p)$$

= (1.602 × 10⁻¹⁹) (2.5 × 10¹³) (3800 + 1800)
= 0.0224 S/cm

(ii) The intrinsic conductivity for silicon,

$$\sigma_i = qn_i (\mu_n + \mu_p)$$

= (1.602 × 10⁻¹⁹) (1.5 × 10¹⁰) (1300 + 500)
= 4.32 × 10⁻⁶ S/cm

2.8 CARRIER CONCENTRATION IN INTRINSIC SEMICONDUCTORS

In an intrinsic semiconductor, the carriers i.e., electrons and holes are generated due to breaking of covalent bonds. Under thermal equilibrium, the number of electrons and holes are equal. At 0K, an intrinsic semiconductor resembles a perfect insulator. To determine the electrical property of a semiconductor, it becomes important to derive the concentration of carriers i.e., electrons 'n' and holes 'p' per unit volume of the material.

2.8.1 Expression for Density of Electrons

Let dn be the number of electrons per m^3 available between energy interval E and E + dE in the conduction band. According to density of states and Fermi distribution function, the number of electrons is given by

2.8 Physics for Electronics Engineering and Information Science

$$dn = N(E) f(E) dE \tag{2.3}$$

As the energy level of the conduction band extends from the bottom of the conduction band E_C to infinity, we have

$$n = \int_{E_C}^{\infty} N(E) f(E) dE$$
(2.4)

where N(E)dE is the density of states in the energy interval *E* and E+dE in the conduction band and f(E) is the probability of electron occupancy.

We know that the density of states i.e., the number of energy states per unit volume within the energy interval E and E+dE (assuming 1 state to 1 electron) is given by

$$N(E)dE = \frac{4\pi}{h^3} (2m)^{3/2} E^{1/2} dE$$

Since the electrons are moving in a periodic potential, its mass 'm' must be replaced by its effective mass m_{e}^{*} .

Hence,

$$N(E)dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} E^{1/2} dE$$

Since *E* starts at the bottom of the conduction band E_C , *E* is replaced by $(E - E_C)$

Therefore,

$$N(E)dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_C)^{1/2} dE$$
(2.5)

From Fermi-Dirac statistics, the probability of an electron occupying an energy state E is given by

$$f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$$

For all possible temperatures, $E - E_F \gg kT$, as $E_C - E_F \ge 4 \text{ kT} \approx 0.1 \text{ eV}$. Hence, neglecting 1 in the denominator, we get

$$f(E) = e^{-\left(\frac{E-E_F}{kT}\right)} = e^{\left(\frac{E_F-E}{kT}\right)}$$
(2.6)

Substitute Eqn. (2.5) and Eqn. (2.6) in Eqn. (2.4), we have

$$n = \int_{E_c}^{\infty} N(E) f(E) dE$$

= $\int_{E_c}^{\infty} \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_C)^{1/2} e^{\left(\frac{E_F - E}{kT}\right)} dE$
= $\frac{4\pi}{h^3} (2m_e^*)^{\frac{3}{2}} e^{\left(\frac{E_F}{kT}\right)} \int_{E_c}^{\infty} (E - E_C)^{1/2} e^{\left(\frac{-E}{kT}\right)} dE$ (2.7)

Substitute $E - E_C = x$, i.e., $E = E_C + x$ and dE = dx. At x = 0, $E = E_C$ and at $x = \infty$, $E = \infty$. Therefore, Eqn. (2.7) becomes

$$n = \frac{4\pi}{h^{3}} (2m_{e}^{*})^{3/2} e^{\left(\frac{E_{F}}{kT}\right)} \int_{0}^{\infty} \sqrt{x} e^{-\left(\frac{E_{C}+x}{kT}\right)} dx = \frac{4\pi}{h^{3}} (2m_{e}^{*})^{3/2} e^{\left(\frac{E_{F}-E_{C}}{kT}\right)} \int_{0}^{\infty} \sqrt{x} e^{-\left(\frac{x}{kT}\right)} dx$$
$$= \frac{4\pi}{h^{3}} \times \frac{\sqrt{\pi}}{2} (2m_{e}^{*})^{3/2} e^{\left(\frac{E_{F}-E_{C}}{kT}\right)} (kT)^{3/2} \qquad [\text{since } \int_{0}^{\infty} \sqrt{x} e^{-\left(\frac{x}{kT}\right)} dx = \frac{\sqrt{\pi}}{2} (kT)^{3/2}]$$
$$= 2 \left(\frac{2\pi m_{e}^{*} kT}{h^{2}}\right)^{3/2} e^{\left(\frac{E_{F}-E_{C}}{kT}\right)}$$
(2.8)
$$\left(\frac{E_{F}-E_{C}}{kT}\right)$$

$$= N_C e^{\left(\frac{L_F - L_C}{kT}\right)}$$
(2.9)

where $N_C = 2 \left(\frac{2\pi m_e^* kT}{h^2} \right)^{3/2}$

Equation (2.9) represents the number of electrons per unit volume in the conduction band of an intrinsic semiconductor.

2.8.2 Expression for Density of Holes

=

Let dp be the number of holes per m³ or vacancies in the energy interval E and E + dE in the valence band. According to density of states and Fermi distribution function, the number of holes is given by

$$dp = N(E)dE\{1 - f(E)\}\$$

$$p = \int_{-\infty}^{E_V} N(E)\{1 - f(E)\}dE$$
(2.10)

where N(E)dE is the density of states in the energy interval *E* and *E* + *dE* in the valence band and 1 - f(E) is the probability of the absence of an electron or the existence of the hole. Here, the lower limit is considered as minus infinity as the probability of existence of holes is zero for the inner shell levels and the value of 1 - f(E) becomes zero anyway.

The density of states in the valence band is

$$N(E)dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} E^{1/2} dE$$

where m_h^* is the effective mass of holes as they move in a periodic potential. Since E_V is the energy of the top of the valence band, *E* is replaced by $(E_V - E)$.

Therefore,

i.e.,

$$N(E)dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_V - E)^{1/2} dE$$
(2.11)

As the top of the valence land E_V is considered as reference, the energy difference is $E_V - E$.

The probability of a hole occupying an energy state E in the valence band i.e., the absence of an electron is given by

$$1 - f(E) = 1 - \left\{\frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}\right\} = \frac{1 + e^{\left(\frac{E - E_F}{kT}\right)} - 1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}} = \frac{e^{\left(\frac{E - E_F}{kT}\right)}}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$$

For all possible temperatures, $E_F - E > kT$ as $E_F - E_V \ge kT \approx 0.1$ eV. Therefore, exponential term in the denominator can be neglected in comparison to 1. Hence,

$$1 - f(E) = e^{\left(\frac{E - E_F}{kT}\right)}$$
(2.12)

Substituting Eqn. (2.11) and Eqn. (2.12) in Eqn. (2.10), we get

$$p = \int_{-\infty}^{E_V} N(E) \{1 - f(E)\} dE = \int_{-\infty}^{E_V} \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_V - E)^{1/2} e^{\left(\frac{E - E_F}{kT}\right)} dE$$
$$= \frac{4\pi}{h^3} (2m_h^*)^{3/2} e^{\left(\frac{-E_F}{kT}\right)} \int_{-\infty}^{E_V} (E_V - E)^{1/2} e^{\left(\frac{E}{kT}\right)} dE$$
(2.13)

Substitute $E_V - E = x$, i.e., $E = E_V - x$ and dE = -dx. At x = 0, $E = E_V$ and at $x = \infty$, $E = -\infty$. Therefore Eq. (2.13) becomes

Therefore, Eqn. (2.13) becomes

$$p = \frac{4\pi}{h^3} (2m_h^*)^{3/2} e^{\left(\frac{-E_F}{kT}\right)} \int_{-\infty}^0 \sqrt{x} e^{\left(\frac{E_V - x}{kT}\right)} (-dx) = \frac{4\pi}{h^3} (2m_h^*)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)} \int_0^\infty \sqrt{x} e^{\left(\frac{-x}{kT}\right)} dx$$
$$= \frac{4\pi}{h^3} \times \frac{\sqrt{\pi}}{2} (2m_h^*)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)} (kT)^{3/2} \qquad [\text{since } \int_0^\infty \sqrt{x} e^{-\left(\frac{x}{kT}\right)} dx = \frac{\sqrt{\pi}}{2} (kT)^{3/2}]$$
$$= 2 \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)}$$
$$p = N_V e^{\left(\frac{E_V - E_F}{kT}\right)}$$
(2.14)

i.e.,

where $N_V = 2 \left(\frac{2\pi m_h^* kT}{h^2} \right)^{3/2}$

Equation (2.14) represents the number of holes per unit volume in the valence band of an intrinsic semiconductor.

2.8.3 Intrinsic Carrier Concentration

If a pure semiconductor is doped with *N*-type impurities, the number of electrons in the conduction band increases above a level and the number of holes in the valence band decreases below a level, which would be available in the intrinsic (pure) semiconductor. Similarly, the addition of *P*-type impurities to a pure semiconductor increases the number of holes in the valence band above a level and decreases the number of electrons in the conduction band below a level, which would have been available in the intrinsic

semiconductor. This is because the rate of recombination increases due to the presence of a large number of free electrons (or holes).

Further, the experimental results state that under thermal equilibrium for any semiconductor, the product of the number of holes and the number of electrons is constant and is independent of the amount of donor and acceptor impurity doping. This relation is known as *mass-action law* and is given by

$$n \cdot p = n_i^2$$

where *n* is the number of free electrons per unit volume, *p* the number of holes per unit volume and n_i is the intrinsic concentration. While considering the conductivity of the doped semiconductors, only the dominant majority charge carriers have to be considered.

The intrinsic concentration is

$$n \cdot p = n_i^2 = 4 \left(\frac{2\pi kT}{h^2}\right)^3 (m_e^* m_h^*)^{3/2} e^{\left(\frac{E_V - E_C}{kT}\right)}$$
$$n_i^2 = 4 \left(\frac{2\pi kT}{h^2}\right)^3 (m_e^* m_h^*)^{3/2} e^{\left(\frac{-E_G}{kT}\right)}$$

i.e.,

where $E_G = E_C - E_V$.

Therefore,

$$n_i = 2\left(\frac{2\pi kT}{h^2}\right)^{3/2} (m_e^* m_h^*)^{3/4} e^{\left(\frac{-E_G}{2kT}\right)}$$
(2.15)

2.8.4 Fermi Level in Intrinsic Semiconductors

In intrinsic semiconductors, the number of electrons is equal to the number of holes i.e., n = p. Equating the electron density with the hole density, we have

$$2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_F - E_C}{kT}\right)} = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)}$$
$$N_C e^{\left(\frac{E_F - E_C}{kT}\right)} = N_V e^{\left(\frac{E_V - E_F}{kT}\right)}$$

Taking logarithm on both sides, we get

$$\ln\frac{N_C}{N_V} = \frac{E_C + E_V - 2E_F}{kT}$$

Therefore, the Fermi level is

$$E_F = \frac{E_C + E_V}{2} - \frac{kT}{2} \ln \frac{N_C}{N_V}$$

At T = 0 K, the Fermi level of intrinsic semiconductor lies in the middle of the valence band and conduction band.

i.e.,
$$E_F = \frac{E_C + E_V}{2}$$

Again equating the electron density with the hole density, we have

$$\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_F - E_C}{kT}\right)} = \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)}$$
$$e^{\left(\frac{2E_F}{kT}\right)} = \left(\frac{m_h^*}{m_e^*}\right)^{3/2} e^{\left(\frac{E_V + E_C}{kT}\right)}$$

Taking logarithm on both sides, we get

$$\frac{2E_F}{kT} = \frac{3}{2}\ln\left(\frac{m_h^*}{m_e^*}\right) + \left(\frac{E_V + E_C}{kT}\right)$$
$$E_F = \frac{3kT}{4}\ln\left(\frac{m_h^*}{m_e^*}\right) + \left(\frac{E_V + E_C}{2}\right)$$

If the effective masses of a free electron and hole are the same i.e., $m_e^* = m_h^*$, then the Fermi level of an intrinsic semiconductor becomes

$$E_F = \left(\frac{E_V + E_C}{2}\right) \tag{2.16}$$

The above equation shows that the Fermi level of intrinsic semiconductor is located exactly halfway between the valence band and conduction band and its position is independent of temperature.



Fig. 2.4 Variation of Fermi level with temperature in intrinsic semiconductor

With increase in temperature, the Fermi level gets displaced to the bottom edge of the conduction band (CB) when $m_h^* > m_e^*$ or to the top edge of the valence band (VB) when $m_h^* < m_e^*$ as seen in Fig. 2.4. The change in Fermi level with the increase in temperature is mainly due to variation of masses of electron and hole with temperature. In practice, the Fermi level in an intrinsic semiconductor is independent of temperature and positioned in the middle of the bandgap when $m_e^* = m_h^*$.

he intrinsic carrier density is 1.5×10^{16} / m³. If the mobilities of electron and hole are 0.13 and 0.05 m^2 / V·s calculate the conductivity. (AU Nov 2003)

Solution Given $\mu_n = 0.13 \text{ m}^2 / \text{V} \cdot \text{s}$ and $n_i = 1.5 \times 10^{16} / \text{m}^3$ at room temperature.

Conductivity,

$$\sigma_i = n_i q = (\mu_n + \mu_p)$$

$$= 1.5 \times 10^{16} \times 1.602 \times 10^{-19} (0.13 + 0.05) = 4.32 \times 10^{-4} (\Omega \cdot m)^{-1}$$

EXAMPLE 2.3

The intrinsic carrier density at room temperature in germanium is 2.37×10^{19} / m³. If the electron and hole mobilities are 0.38 and 0.18 m²/V · s respectively, calculate the resistivity. (AU Oct 2002, Dec 2010)

Solution Given $\mu_n = 0.38 \text{ m}^2/\text{V} \cdot \text{s}$, $\mu_p = 0.18 \text{ m}^2/\text{V} \cdot \text{s}$ and $n_i = 2.37 \times 10^{19} / \text{m}^3$ at room temperature.

Conductivity,

$$= 2.37 \times 10^{19} \times 1.602 \times 10^{-19} (0.38 + 0.18) = 2.12 (\Omega \cdot m)^{-1}$$

Resistivity,

$$\rho_i = \frac{1}{\sigma_i} = \frac{1}{2.12} = 0.47 \,\Omega \cdot \mathrm{m}$$

 $\sigma_i = n_i q(\mu_n + \mu_p)$

EXAMPLE 2.4

The electron mobility and hole mobility in Si are $0.135 \text{ m}^2/\text{V} \cdot \text{s}$ and $0.048 \text{ m}^2/\text{V} \cdot \text{s}$ respectively, at room temperature. If the carrier concentration is $1.5 \times 10^{16} / \text{m}^3$. Calculate the resistivity of Si at room temperature. (AU June 2009)

Solution Given $\mu_n = 0.135 \text{ m}^2/\text{V} \cdot \text{s}$, $\mu_n = 0.048 \text{ m}^2/\text{V} \cdot \text{s}$ and $n_i = 1.5 \times 10^{16} / \text{m}^3$.

Conductivity,

$$\sigma_i = n_i q(\mu_n + \mu_p)$$

= 1.5 × 10¹⁶ × 1.602 × 10⁻¹⁹ (0.135 + 0.048) = 0.439 × 10⁻³ (Ω · m)⁻¹
$$\rho_i = \frac{1}{\sigma_i} = \frac{1}{0.439 \times 10^{-3}} = 2.28 \times 10^3 \,\Omega \cdot m$$

Resistivity,

EXAMPLE 2.5

Find the resistance of an intrinsic Ge rod 1 mm long, 1 mm wide and 1 mm thick at 300 K. The intrinsic carrier density is 2.5×10^{19} / m³ at 300 K and the mobility of electron and hole are 0.39 and 0.19 m²/V · s. (AU April 2003, June 2012)

Solution Given l = w = t = 1 mm, $\mu_n = 0.39 \text{ m}^2/\text{V} \cdot \text{s}$, $\mu_p = 0.19 \text{ m}^2/\text{V} \cdot \text{s}$ and $n_i = 2.5 \times 10^{19} / \text{m}^3$.

Conductivity, $\sigma_i = n_i q(\mu_p + \mu_n)$

Resistance,

$$R = \frac{l}{\sigma A} = \frac{l}{n_i q(\mu_p + \mu_n)A}$$

$$=\frac{1\times10^{-3}}{2.5\times10^{19}\times1.602\times10^{-19}\times(0.39+0.19)\times1\times10^{-6}}=431\,\Omega$$

For an intrinsic Semiconductor with a band gap of 0.7 eV, determine the position of E_F at T = 300 K if $m_h^* = 6m_{e^*}^*$ (AU Nov 2003)

Solution Given $E_G = 0.7 \text{ eV}$, and $m_h^* = 6m_e^*$.

Fermi level,

$$E_F = \frac{E_G}{2} + \frac{3kT}{4} \ln\left(\frac{m_h^*}{m_e^*}\right) = \frac{E_G}{2} + \frac{3kT}{4} \ln(6)$$

where Boltzmann constant, $k = 1.38 \times 10^{-23}$ J/K or 8.61 × 10⁻⁵ eV/K

Therefore,

$$E_F = \frac{0.7}{2} + \frac{3 \times 8.61 \times 10^{-5} \times 300 \times 1.79}{4} = 0.385 \text{ eV}$$

EXAMPLE 2.7

Find the resistance of an intrinsic germanium rod 1.5 cm long, 1 m wide and 2 m thick at 300 K. For Germanium $n_i = 2.3 \times 10^{19} / \text{m}^3$, $\mu_n = 0.35 \text{ m}^2 / \text{V} \cdot \text{s}$ and $\mu_p = 0.17 \text{ m}^2 / \text{V} \cdot \text{s}$ at 300 K

Solution Given l = 1.5 cm, A = 2 mm², $\mu_n = 0.35$ m²/V · s, $\mu_p = 0.17$ m²/V · s and $n_i = 2.3 \times 10^{19}$ / m³.

We know that conductivity, $\sigma = n_i q(\mu_p + \mu_n)$

Resistance,

$$R = \frac{l}{\sigma A} = \frac{l}{n_i q(\mu_p + \mu_n)A}$$

$$=\frac{1.5\times10^{-2}}{2.3\times10^{19}\times1.602\times10^{-19}\times(0.35+0.17)\times2\times10^{-6}}=3.919\times10^{3}\ \Omega$$

EXAMPLE 2.8

Calculate the current produced in a small silicon plate of area 1 cm² and thickness 0.5 m when a potential difference of 2 V is applied across the faces. Given the concentration of free electrons in silicon is 1.45×10^{10} / m³ and the mobilities of electron and holes are 0.15 m²/V · s and 0.445 m²/V · s respectively.

Solution Given t = 1.5 mm, A = 1 cm², $\mu_n = 0.15$ m²/V · s, $\mu_p = 0.45$ m²/V · s, $n_i = 1.45 \times 10^{10}$ / m³ and V = 2 V.

We know that conductivity, $\sigma = \frac{J}{E} = \frac{I/A}{V/l} = n_i q(\mu_p + \mu_n)$

=

Therefore, the current is I

$$I = n_i q(\mu_n + \mu_p) \times \frac{A \times V}{l}$$

= $\frac{1.45 \times 10^{10} \times 1.602 \times 10^{-19} \times 0.6 \times 1 \times 10^{-4} \times 2}{05 \times 10^{-3}} = 55.6 \text{ nA}$

The resistivity of intrinsic Si is $2.3 \times 10^3 \Omega \cdot m$ at 300 K. Calculate its resistivity at 100°C. Assume $E_G = 1 \text{ eV}$ and $k = 1.38 \times 10^{-23} \text{ J/K}.$

Solution Given $\rho_1 = 2.3 \times 10^3 \,\Omega \cdot \text{m}$ at $T_1 = 300 \,\text{K}$, $E_G = 1 \,\text{eV}$ and $k = 1.38 \times 10^{-23} \,\text{J/K} = 8.61 \times 10^{-5} \,\text{eV/K}$. The resistivity of the intrinsic semiconductor (Si) is $\rho_i = \rho_0 e^{E_G/2kT}$

At
$$T_1 = 300$$
 K, $\rho_1 = \rho_0 e^{\frac{E_G}{2k(300)}}$

At $T_2 = 373$ K, $\rho_2 = \rho_0 e^{\frac{E_G}{2k(373)}}$

$$\rho_2 = \rho_1 \times 0.0225 = 2.3 \times 10^3 \times 0.0225 = 51.75 \ \Omega \cdot m$$

EXAMPLE 2.10

The electrical conductivity of Ge at 20°C is 2 $(\Omega \cdot m)^{-1}$. What is its conductivity at 40°C? Bandgap of Ge is 0.72 eV.

Solution Given $\sigma_1 = 2 (\Omega \cdot m)^{-1}$ at $T_1 = 20^{\circ}$ C, $E_G = 0.72 \text{ eV}$ and $k = 1.38 \times 10^{-23} \text{ J/K} = 8.61 \times 10^{-5} \text{ eV/K}$. The conductivity of the intrinsic semiconductor (Ge) is $\sigma_i = \sigma_0 e^{-E_G/2kT}$

At
$$T_1 = 293$$
 K, $\sigma_1 = \sigma_0 e^{-\frac{E_G}{2k(293)}}$

At
$$T_1 = 293$$
 K, $\sigma_1 = \sigma_0 e^{-\frac{E_G}{2k(313)}}$
At $T_2 = 313$ K, $\sigma_2 = \sigma_0 e^{-\frac{E_G}{2k(313)}}$

$$\frac{\sigma_2}{\sigma_1} = \frac{e^{\left[\frac{0.72}{2 \times 8.61 \times 10^{-5} \times 293}\right]}}{e^{\left[\frac{0.72}{2 \times 8.61 \times 10^{-5} \times 313}\right]}} = \frac{e^{14.27}}{e^{13.36}} = 2.48$$

Hence,

$$\sigma_2 = \sigma_1 \times 2.48 = 2 \times 2.48 = 4.96 (\Omega \cdot m)^{-1}$$

EXAMPLE 2.11

Assuming that the number of electrons near the top of the valence band available for thermal excitation is 5×10^{25} / m³ and the intrinsic carrier density is 2.5×10^{19} / m³, calculate the energy gap of germanium at room temperature.

$$\frac{\rho_2}{\rho_1} = \frac{e^{\left[\frac{1}{2 \times 8.61 \times 10^{-5} \times 373}\right]}}{e^{\left[\frac{1}{2 \times 8.61 \times 10^{-5} \times 300}\right]}} = \frac{e^{15.57}}{e^{19.36}} = 0.0225$$

Solution Given the density of electrons, $N = 5 \times 10^{25} / \text{m}^3$ and the intrinsic carrier density, $n_i = 2.5 \times 10^{10} / \text{m}^3$. The fraction of electron in conduction band, $\frac{n_i}{N} = e^{-E_G/2kT}$

Therefore,

$$E_G = -2kT \ln\left(\frac{n_i}{N}\right) = -2 \times 8.61 \times 10^{-5} \times 300 \times \ln\left[\frac{2.5 \times 10^{19}}{5 \times 10^{25}}\right]$$

 $= 2 \times 0.026 \times 14.509 = 0.75 \text{ eV}.$

EXAMPLE 2.12

Estimate the fraction of electrons in conduction band at room temperature in Ge with $E_G = 72$ eV and in diamond with $E_G = 5.6$ eV.

Solution The fraction of electrons in conduction band, $\frac{n_i}{N} = e^{-E_G/2kT}$ where kT = 0.026 eV.

(i) The fraction of electrons in conduction band in E_G with $E_G = 0.72$ eV is

$$\frac{n_i}{N} = e^{\left[-\frac{0.72}{2 \times 0.026}\right]} = 9.7 \times 10^{-7}$$

(ii) The fraction of electrons in conduction band in diamond with $E_G = 5.6$ eV is

$$\frac{n_i}{N} = e^{\left[-\frac{5.6}{2 \times 0.026}\right]} = 1.7 \times 10^{-47}$$

EXAMPLE 2.13

Calculate the number of electron-hole pairs per cubic metre in intrinsic Si at 310 K. the value of its forbidden energy gap is 1.12 eV. The number of electron-hole pairs in pure Ge per m³ is 2.5×10^{19} at 300 K and the value of E_G for Ge is 0.72 eV. ($k = 1.38 \times 10^{-23}$ J/K).

Solution Given $n_{Ge} = 2.5 \times 10^{19}$ / m³ at $T_1 = 300$ K, $E_{G1} = 0.72$ eV for Ge and $E_{G2} = 1.12$ eV at $T_2 = 310$ K for Si.

Intrinsic concentration for Ge, $n_{Ge} = 2\left[\frac{2\pi kT_1}{h^2}\right]^{3/2} \times (m_e^* m_h^*)^{3/4} \times e^{-E_{G1}/2kT_1}$

where
$$A = 2 \left[\frac{2\pi k}{h^2} \right]^{3/2} \times (m_e^* m_h^*)^{\frac{3}{4}}$$

Hence,

$$2.5 \times 10^{19} = A(300)^{3/2} \times e^{\left(\frac{-1.602 \times 10^{-19} \times 0.72}{2 \times 1.38 \times 10^{-23} \times 300}\right)} = A(300)^{3/2} \times 9.1 \times 10^{-7}$$
(1)

Intrinsic concentration for Si, $n_{Si} = 2 \left[\frac{2\pi kT_2}{h^2} \right]^{3/2} \times \left(m_e^* m_h^* \right)^{\frac{3}{4}} \times e^{-E_{G2}/2kT_2}$

-212

i.e.,
$$n_{Si} = A(310)^{3/2} \times e^{\left(\frac{-1.602 \times 10^{-19} \times 1.12}{2 \times 1.38 \times 10^{-23} \times 310}\right)} = A(310)^{3/2} \times 8.05 \times 10^{-10}$$
 (2)

Dividing Eqn. (2) by Eqn. (1), we get

$$\frac{n_{Si}}{2.5 \times 10^{19}} = \left(\frac{310}{300}\right)^{3/2} \times \frac{8.05 \times 10^{-10}}{9.1 \times 10^{-7}} = 0.93 \times 10^{-3}$$

Therefore,

$$n_s = 2.325 \times 10^{-16} / \text{m}^3$$

EXAMPLE 2.14

In an intrinsic semiconductor, the effective mass of the electron is 0.07 m_0 and that of the hole is $0.4 m_0$ where m_0 is the rest mass of the electron. Calculate the intrinsic concentration of charge carriers at 300 K. Given $E_G = 0.7$ eV.

Solution Given $E_G = 0.7 \text{eV}, m_e^* = 0.07 m_0, m_h^* = 0.4 m_0$ and assume $m_0 = 9.11 \times 10^{-31} \text{ kg},$

Intrinsic concentration, $n_i = 2 \left[\frac{2\pi kT}{h^2} \right]^{3/2} \times (m_e^* m_h^*)^{\frac{3}{4}} \times e^{-E_G/2kT}$

Here,

$$2\left(\frac{2\pi m_0 kT}{h^2}\right)^{3/2} = 2\left[\frac{2\times 3.14\times 9.11\times 10^{-31}\times 1.38\times 10^{-23}\times 300}{(6.626\times 10^{-34})^2}\right]^{3/2} = 2.5\times 10^{25}$$
$$(m_e^* m_h^*)^{3/4} = (0.07\times 0.4)^{3/4} = 0.068$$

and

$$E_G/2kT = e^{\left(-\frac{0.7 \times 1.602 \times 10^{-2}}{2 \times 1.38 \times 10^{-23} \times 300}\right)} = e^{-13.53} = 1.33 \times 10^{-6}$$

Therefore,

$$n_i = 2.5 \times 10^{25} \times 0.068 \times 1.33 \times 10^{-6} = 2.26 \times 10^{18} / \text{m}^3$$

EXAMPLE 2.15

e

Compute the concentration of intrinsic charge carriers in a germanium crystal at 300 K. Given that $E_G = 0.72$ eV and assume $m_e^* = m$.

Solution Given $E_G = 0.72$ eV at T = 300 K for germanium and assume $m_e^* = 9.11 \times 10^{-31}$ kg.

Intrinsic charge carrier concentration, $n_i = 2 \left[\frac{2\pi m_e^* kT}{h^2} \right]^{3/2} \times e^{-E_G/2kT}$

$$2\left(\frac{2\pi mkT}{h^2}\right)^{3/2} = 2\left[\frac{2\times3.14\times9.11\times10^{-31}\times1.38\times10^{-23}\times300}{(6.626\times10^{-34})^2}\right]^{3/2} = 2.5\times10^{25}$$

Here,

$$e^{-E_G/2kT} = e^{\left(-\frac{0.72 \times 1.602 \times 10^{-19}}{2 \times 1.38 \times 10^{-23} \times 300}\right)} = e^{-13.846} = 0.97 \times 10^{-6}$$

and

Therefore,

$$n_i = 2.5 \times 10^{25} \times 0.97 \times 10^{-6} = 2.425 \times 10^{19} / \text{m}^3$$

EXAMPLE 2.16

The band gap energy of an intrinsic semiconductor in silicon is 1.1 eV. Calculate the concentration of intrinsic charge carriers at 300 K(Assume $m_e^* = m$).

Solution Given $E_G = 1.1$ eV at T = 300 K for silicon and assume $m_e^* = m = 9.11 \times 10^{-31}$ kg.

Intrinsic charge carrier concentration, $n_i = 2 \left[\frac{2\pi m_e^* kT}{h^2} \right]^{3/2} \times e^{-E_G/2kT}$

Here,

$$2\left(\frac{2\pi mkT}{h^2}\right)^{3/2} = 2\left[\frac{2\times3.14\times9.11\times10^{-31}\times1.38\times10^{-23}\times300}{(6.626\times10^{-34})^2}\right]^{3/2} = 2.5\times10^{25}$$

$$e^{-E_G/2kT} = e^{\left(\frac{-1.11 \times 1.602 \times 10^{-19}}{2 \times 1.38 \times 10^{-23} \times 300}\right)} = e^{-21.44} = 4.88 \times 10^{-10}$$

Therefore,

$$n_i = 2.5 \times 10^{25} \times 4.88 \times 10^{-10} = 1.22 \times 10^{16} / \text{m}^3$$

EXAMPLE 2.17

The energy gap of Si is 1.1 eV. Its electron and hole mobilities at room temperature are 0.48 and $0.013 \text{ m}^2/\text{V} \cdot \text{s}$ respectively. Find the conductivity σ .

Solution Given $E_G = 1.1 \text{ eV}$ at room temperature for Si, $\mu_n = 0.48 \text{ m}^2/\text{V} \cdot \text{s}$ and $\mu_p = 0.013 \text{ m}^2/\text{V} \cdot \text{s}$ Intrinsic carrier concentration, $n_i = 1.22 \times 10^{16} / \text{m}^3$ (Refer previous example)

Conductivity,

 $\sigma_i = n_i q(\mu_n + \mu_p)$ = 1.22×10¹⁶×1.602×10⁻¹⁹(0.48 + 0.013) = 0.96×10⁻³ (Ω · m)⁻¹

EXAMPLE 2.18

The intrinsic resistivity of Ge at 300 K is 47 Ω ·cm. What is the intrinsic carrier concentration? Also calculate the drift velocity of holes and electrons for an electric field. Assume E = 100 V/cm, $\mu_n = 0.39$ m²/V·s and $\mu_n = 0.19$ m²/V·s

Solution Given $\rho = 47 \ \Omega$. cm, E = 100 V/cm, $\mu_n = 0.39 \text{ m}^2 / \text{V} \cdot \text{s}$ and $\mu_p = 0.19 \text{ m}^2 / \text{V} \cdot \text{s}$

We know that $\sigma_i = n_i q(\mu_n + \mu_p)$

Hence, the intrinsic carrier concentration is

$$n_i = \frac{\sigma_i}{q(\mu_n + \mu_p)} = \frac{1}{\rho_i q(\mu_n + \mu_p)}$$

$$=\frac{1}{47 \times 10^{-2} \times 1.602 \times 10^{-19} \times (0.39 + 0.19)} = 2.293 \times 10^{19} / \text{m}^3$$

Drift velocity of holes, $v_p = \mu_p E = 0.19 \times 10^4 = 1900$ m/s

Drift velocity of electrons, $v_n = \mu_n E = 0.39 \times 10^4 = 3900$ m/s

EXAMPLE 2.19

Determine the position of Fermi level in silicon semiconductor at 300 K. Given that the band gap is $1.2 \text{ eV}, m_e^* = 0.12 \text{ m}$ and $m_h^* = 0.28 \text{ m}$.

Solution Given $E_G = 1.12 \text{ eV}$, $m_e^* = 0.12 \text{ m}$ and $m_h^* = 0.28 \text{ m}$. Here, m is the mass of an electron.

We know that the Fermi level, $E_F = \frac{E_G}{2} + \frac{3 kT}{4} \ln \left(\frac{m_h^*}{m_e^*}\right)$

$$E_F = \frac{1.12}{2} + \frac{3 \times 8.61 \times 10^{-5} \times 300}{4} \ln\left(\frac{0.28}{0.12}\right)$$

 $= 0.56 + (0.0194) \ln 2.333 = 0.576 \text{ eV}$

Hence, the Fermi level is 0.016 eV above the centre of the forbidden gap. In other words, it is at 0.576 eV from the top of the valence band.

EXAMPLE 2.20

The resistivity of an intrinsic semiconductor is 4.5 Ω ·m at 20°C and 2 Ω ·m at 20°C Find the energy gap.

Solution Given $\rho_1 = 4.5 \ \Omega \cdot m$ at $T_1 = 20^{\circ}C = 293 \text{ K}$ and $\rho_2 = 2 \ \Omega \cdot m$ at $T_2 = 32^{\circ}C = 305 \text{ K}$.

We know that the energy bandgap, $E_G = 2k \left(\frac{dy}{dx}\right)$

$$E_G = 2k \left(\frac{\log \rho_2 - \log \rho_1}{\left(\frac{1}{T_1} - \frac{1}{T_2}\right)} \right) = 2 \times 8.61 \times 10^{-5} \times \left(\frac{\log 2 - \log 4.5}{\left(\frac{1}{293}\right) - \left(\frac{1}{305}\right)} \right)$$
$$= 17.22 \times 10^{-5} \times \left(\frac{0.6532 - 0.3010}{0.134 \times 10^{-3}} \right) = 0.45 \text{ eV}$$

2.9 PROPERTIES OF INTRINSIC SEMICONDUCTORS

Table 2.2 gives the important properties of intrinsic silicon and germanium at room temperature (300 K).

	Si	Ge	GaAs
Atomic number	14	32	_
Atomic weight	28.09	72.59	_
Atomic density, (m ⁻³)	5.02×10^{28}	4.42×10^{28}	_
Lattice constant, a (nm)	0.543	0.565	0.563
Relative permittivity, ε_r	11.8	16.0	13.5
Density, g/cm ³	2.33	5.32	_
Energy gap, E_G (eV)	1.08	0.66	1.58
Electron mobility, μ_n (m ² /V–s)	0.13	0.38	0.85
Hole mobility, $\mu_p (\text{m}^2/\text{V}-\text{s})$	0.05	0.18	0.04
Intrinsic concentration, n_i (m ⁻³)	1.38×10^{16}	2.5×10^{19}	9×10^{12}
Electron diffusion constant, $D_n (m^2/S) = \mu_n V_T$	0.0034	0.0099	0.020
Hole diffusion constant, $D_P (m^2/S) = \mu_p V_T$	0.0013	0.0047	_
Density of states at conduction band edge, N_C (m ⁻³)	2.8×10^{25}	1.0×10^{25}	4.7×10^{23}
Density of states at valence band edge, N_V (m ⁻³)	1.0×10^{25}	6.0×10^{24}	7.0×10^{24}
Intrinsic resistivity, Ω-cm	23×10^4	45	_
Melting point	1420	936	1250

 Table 2.2
 Properties of some common semiconductors at room temperature

2.10 EXTRINSIC SEMICONDUCTORS WITH ENERGY BAND DIAGRAM

In pure semiconductors, the carrier concentration is very small at room temperature. In Silicon (Si) and Gallium Arsenide (GaAs), these are of the order of 1.5×10^{10} per cm³ and 1.1×10^{6} electrons per cm³ at 298 K, which are quite lesser than the order of 10^{28} electrons per cm³ in metals. Hence, the electrical resistivity is very high in pure semiconductors. The resistivity can be decreased by adding impurities to a pure semiconductor. This process is called *doping* and the impurity atoms are called *dopants*. Even a 0.0001% level addition of impurity atoms causes a large increase in the conductivity of a semiconductor.

Semiconductors (Si or Ge) are typically doped with elements such as Boron, Arsenic and Phosphorous to change and enhance their electrical properties. By doping, a crystal can be modified so that it has a predominance of either electrons or holes. Hence, there are two types of doped semiconductors, N-type (mostly electrons) and P-type (mostly holes). When a crystal is doped such that the equilibrium carrier concentrations n and p are different from the intrinsic carrier concentration n_i , the material is said to be *extrinsic*.

The amount of impurities doped determines the increase in conductivity at room temperature. For example, by adding a donor impurity of 1 part in 10^8 , the conductivity of germanium at 30° C increases twelve times. Adding 1 Boron (B) atom for every 10^5 Germanium (Ge) atoms increases the conductivity by nearly ten thousand times. Thus, the donor or acceptor concentration determines the increase in conductivity. When

impurities or lattice defects are introduced, additional levels are created in the energy bands structure, usually within the band gap.

The valence and conduction bands of Silicon (Si) with additional impurity energy levels within the energy gap are shown in Fig. 2.5. It is noted that the values indicate the change in bandgap due to impurities at T = 0 K. Here, the donor impurities are from Group V elements like Antimony (Sb), Phosphorus (P) and Arsenic (As) for *N*-type semiconductors. The acceptor impurities are from Group III elements like Boron (B), Aluminium (Al) and Gallium (Ga) for *P*-type semiconductors.

2.10.1 N-Type Semiconductors

In the covalent bonding model, donor atoms can be visualized as shown in the Fig. 2.6. When pentavalent elements such as Sb, P or As atom from group V is added with a pure semiconductor like Ge or Si, the lattice structure has four necessary valence electrons to complete the covalent bonds with the neighboring Ge or Si atoms, plus one extra electron. This fifth electron does not fit into the bonding structure of

the lattice and is therefore, loosely bound to the Antimony (Sb) atom at T = 0 K as shown in Fig. 2.6(a). At T > 0 K, a small amount of thermal energy enables this extra electron to overcome its covalent binding to the impurity atom and move into the conduction band and is free to participate in current conduction. Now, a positively ionized Sb atom is left in the structure due to its donation of a free electron. Since the Sb atoms donate electrons, they are called *donors*.



Fig. 2.6 Covalent Bond and Energy Band diagram for an N-type semiconductor at (a) T = 0 K and (b) T > 0 K



The covalent bond and energy band diagram for an *N*-type semiconductor are shown in Fig. 2.6 for (a) T = 0 K and (b) T > 0 K. An impurity from group V introduces an energy level very near the conduction band in Ge or Si. This level is filled with electrons at 0 K, and very little thermal energy is required to excite these electrons to the conduction band. Thus, at about 50 to100 K nearly all of the electrons in the impurity level are "donated" to the conduction band. Such an impurity level is called a *donor* level, and the pentavalent impurities in Ge or Si are called donor impurities. When the temperature is too low for the intrinsic electronhole pair concentration to be appreciable, the density of the electrons donated will be significantly high. Thus, the electrons will be the majority carriers at room temperature and hence, these materials are called *N*-type semiconductors.

2.10.2 P-Type Semiconductors

When trivalent elements such as B, Al, Ga or Indium from group III which have three electrons in their outer most orbits, are added to the pure semiconductor like Ge or Si, all the three electrons of boron are engaged in covalent bonding with the three neighboring Ge or Si atoms. Boron (B) now needs one more electron to complete its bond. This electron may be supplied by Ge or Si, thereby creating a vacant electron site or hole on the semiconductor atom. As Boron accepts one extra electron, the energy level of this impurity atom is called acceptor level and this acceptor level lies just above the valence band. These types of trivalent impurities from group III are called acceptor impurities and the semiconductors doped with these are called *P*-type semiconductors. The covalent bond and energy band diagram for a *P*-type semiconductor are shown in Fig. 2.7 for (a) T = 0 K and (b) T > 0 K.



Fig. 2.7 Covalent Bond and Energy Band diagram for a P-type semiconductor at (a) T = 0 K and (b) T > 0 K

At very low temperatures, electrons from the valence band fill the acceptor levels and hence they become negatively ionized. Hence, holes are created in the valence band. As more and more acceptors are ionized, more holes are created. Since the holes exceed the number of electrons, they form the majority carriers and electrons form the minority carriers in *P-type semiconductors*.

2.11 CARRIER CONCENTRATION IN *N*-TYPE AND *P*-TYPE SEMICONDUCTORS

The distribution of electrons and holes in the semiconductor changes due to the addition of donor or acceptor impurity atoms (dopants) to a pure semiconductor. As the dopant atoms are added, the Fermi energy will change due to its relation with the distribution function. When donor impurity atoms are added, the semiconductor becomes N-type and the density of electrons in N-type semiconductor will be greater than the density of holes. When acceptor impurity atoms are added, the semiconductor becomes P-type and the density of holes in P-type semiconductor will be greater than the density of holes in P-type semiconductor will be greater than the density of electrons. Hence, there will be a change in Fermi energy level in a semiconductor due to change in the electron and hole concentrations. Further, the Fermi level changes due to addition of donor or acceptor impurity atoms.

2.11.1 N-type Semiconductors

The energy level diagram of an *N*-type semiconductor is shown in Fig. 2.6. Let N_D be the donor concentration i.e., the number of donor atoms per unit volume of the material and E_D be the donor energy level. At very low temperatures, all donor levels are filled with electrons. As the temperature increases, more and more donor atoms get ionized. Hence, the electron density in the conduction band increases.

The density of electrons 'n' in the conduction band is

$$n = 2 \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_F - E_C}{kT}\right)} = N_C e^{\left(\frac{E_F - E_C}{kT}\right)}$$
(2.17)

where N_C is the effective density of states in the conduction band, m_e^* is the effective mass of an electron, k is the Boltzmann constant and h is the Planck's constant.

The parameter N_C in Eqn. (2.17) is given by

$$N_C = 2 \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2}$$

For an *N*-type semiconductor, $n = N_D = N_C e^{\left(\frac{E_F - E_C}{kT}\right)}$ where N_D is the concentration of donor atoms.

Therefore, position of the Fermi level with reference to the edge of the conduction band is

$$E_C - E_F = kT \ln \frac{N_C}{N_D} \tag{2.18}$$

Fermi level at 0 K

Consider the probability of absence of an electron at level E_D as shown in Fig. 2.8. Since E_F lies above the donor level, the density of ionized donors is

$$\begin{split} N_D[1-f(E_D)] &= N_d \left[1 - \frac{1}{1 + e^{\left(\frac{E_D - E_F}{kT}\right)}} \right] = N_D \left[\frac{1 + e^{\left(\frac{E_D - E_F}{kT}\right)}}{1 + e^{\left(\frac{E_D - E_F}{kT}\right)}} \right] \\ &\cong N_D e^{\left(\frac{E_D - E_F}{kT}\right)} \quad \text{(for } E_F - E_D \ge kT) \end{split}$$

At very low temperatures, when electron-hole pairs are not generated due to breaking of covalent bonds, the number of electrons in the conduction band must be equal to the number of ionized donors.

Therefore,

$$2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{\frac{3}{2}} e^{\left(\frac{E_F - E_C}{kT}\right)} = N_D e^{\left(\frac{E_D - E_F}{kT}\right)}$$

Taking logarithm on both sides and rearranging, we get

$$\left(\frac{E_F - E_C}{kT}\right) - \left(\frac{E_D - E_F}{kT}\right) = \ln N_D - \ln 2 \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2}$$
$$2E_F - (E_D + E_C) = kT \ln \frac{N_D}{2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2}}$$



Fig. 2.8 Fermi level for N-type semiconductor at low temperature

Hence, the Fermi level is

$$E_F = \frac{(E_D + E_C)}{2} + \frac{kT}{2} \ln \frac{N_D}{2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2}}$$
(2.19)

 $E_F = \frac{(E_D + E_C)}{2}$ (2.20)The above equation shows that Fermi level E_F lies at the middle of the donor level and the bottom of the conduction band at 0 K.

Density of electrons in the conduction band

The electron density i.e., the number of electrons per m³ in the conduction band is given by

$$n = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_F - E_C}{kT}\right)}$$

Here,

$$e^{\left(\frac{E_{E}-E_{C}}{kT}\right)} = \exp\left\{\frac{\left\{\frac{\left(E_{D}+E_{C}\right)}{2} + \frac{kT}{2}\ln\frac{N_{D}}{2\left(\frac{2\pi m_{e}^{*}kT}{h^{2}}\right)^{3/2}}\right\} - E_{C}}{kT}\right\}$$

$$= \exp\left\{\frac{\left(E_{D}+E_{C}\right)}{2kT} + \ln\frac{\left(N_{D}\right)^{1/2}}{\left[2\left(\frac{2\pi m_{e}^{*}kT}{h^{2}}\right)^{3/2}\right]^{1/2}} - \frac{E_{C}}{kT}\right\}$$
(since $\frac{1}{2}\ln x = \ln x^{1/2}$)
$$= \exp\left\{\frac{\left(E_{D}-E_{C}\right)}{2kT} + \ln\frac{\left(N_{D}\right)^{1/2}}{\left[2\left(\frac{2\pi m_{e}^{*}kT}{h^{2}}\right)^{3/2}\right]^{1/2}}\right\}$$

$$e^{\left(\frac{E_{E}-E_{C}}{kT}\right)} = \frac{\left(N_{D}\right)^{1/2}}{e^{\left(\frac{E_{D}-E_{C}}{2kT}\right)}} e^{\left(\frac{E_{D}-E_{C}}{2kT}\right)}}$$
(since $e^{(a+b)} = e^{a} + e^{b}$)

i.e.,

 $\frac{(E)}{\left[2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2}\right]^{1/2}} e^{\frac{(E_D - E_C)}{2kT}} \qquad (\text{since } e^{(a+b)} = e^a + e^b)$

Now, the electron density is

$$n = 2 \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_F - E_C}{kT}\right)}$$
$$= 2 \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \times \frac{(N_D)^{1/2}}{\left[2^{1/2} \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/4}\right]} e^{\frac{(E_D - E_C)}{2kT}}$$
$$= (2N_D)^{1/2} \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/4} e^{\frac{(E_D - E_C)}{2kT}}$$
(2.21)

The above equation signifies that at moderately low temperatures, the electron density in the conduction band is proportional to the square root of donor concentration N_D .

2.11.2 P-type Semiconductors

The energy level diagram of a *P*-type semiconductor is shown in Fig. 2.7. Let N_A be the acceptor concentration i.e., the number of acceptor atoms per unit volume of the material and E_A be the acceptor energy level. At very low temperatures, all acceptor levels are empty. With increase in temperature, more and more acceptor atoms get ionized i.e., the electrons move from valence band and occupy the vacant sites in the acceptor energy level, thereby, leaving holes in the valence band.

The density of holes 'p' in the valence band is

$$p = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)} = N_V e^{\left(\frac{E_V - E_F}{kT}\right)}$$
(2.22)

where N_V is the effective density of states in the valence band, m_h^* is the effective mass of a hole, k is the Boltzmann constant and h is the Planck's constant.

The parameter N_V in Eqn. (2.22) is given by

$$N_V = 2 \left(\frac{2\pi m_h^* kT}{h^2} \right)^{3/2}$$

For a *P*-type semiconductor, $p = N_A = N_V e^{\left(\frac{E_V - E_F}{kT}\right)}$ where N_A is the concentration of acceptor atoms. Therefore, position of the Fermi level with reference to the edge of the valence band is

$$E_F - E_V = kT \ln \frac{N_V}{N_A} \tag{2.23}$$

Fermi level at 0 K

Since E_F lies below the acceptor level as shown in Fig. 2.9, the density of ionized acceptors is

$$N_A f(E_A) \approx N_A e^{\left(\frac{E_F - E_A}{kT}\right)}$$
$$f(E_A) = \frac{1}{1 + e^{\left(\frac{E_A - E_F}{kT}\right)}}$$

Here,

For $E_A - E_F > kT$, the exponential term in the above equation is very large and neglecting '1', we get

 $N_A f(E_A) = N_A e^{\left(\frac{E_F - E_A}{kT}\right)}$

$$f(E_A) = e^{-\left(\frac{E_A - E_F}{kT}\right)}$$

Therefore,





Fig. 2.9 Fermi level for P-type semiconductor at low temperature

i.e.,
$$2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)} = N_A e^{\left(\frac{E_F - E_A}{kT}\right)}$$
$$e^{\left(\frac{E_V + E_A - 2E_F}{kT}\right)} = \frac{N_A}{2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}}$$

Taking logarithm on both sides, we get

$$\left(\frac{E_V + E_A - 2E_F}{kT}\right) = \ln \frac{N_A}{2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}}$$

Hence,

$$E_F = \frac{E_V + E_A}{2} - \frac{kT}{2} \ln \frac{N_A}{2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}}$$
(2.24)

At 0 K,
$$E_F = \frac{E_V + E_A}{2}$$
 (2.25)

The above equation shows that Fermi level E_F lies at the middle of the acceptor level and the top of the valence band at 0 K.

Density of holes in the valence band

The hole density i.e., the number of holes per m³ in the valence band is given by

$$p = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_V - E_F}{kT}\right)}$$

Using Eqn. (2.24), we can write

$$E_{V} - E_{F} = E_{V} - \frac{E_{V} + E_{A}}{2} + \frac{kT}{2} \ln \frac{N_{A}}{2\left(\frac{2\pi m_{h}^{*} kT}{h^{2}}\right)^{3/2}}$$

$$\frac{E_V - E_F}{kT} = \frac{E_V - E_A}{2kT} + \frac{1}{2} \ln \frac{N_A}{2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}}$$

Taking exponential on both side of the above equation, we get

ſ

$$e^{\left(\frac{E_V - E_F}{kT}\right)} = \exp\left\{\frac{E_V - E_A}{2kT} + \frac{1}{2}\ln\frac{N_A}{2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}}\right\}$$
$$= \exp\left\{\frac{\left(\frac{E_V - E_A}{2kT}\right) + \ln\frac{N_A^{1/2}}{\left[2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}\right]^{1/2}}\right\}$$
(since $\frac{1}{2}\ln x = \ln x^{1/2}$)
$$= e^{\left(\frac{E_V - E_A}{2kT}\right)} \times \frac{N_A^{1/2}}{\left[2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}\right]^{1/2}}$$
(since $e^{(a+b)} = e^a + e^b$)

Now, the hole density is

$$p = 2 \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \left[e^{\left(\frac{E_V - E_A}{2kT}\right)} \right] \frac{N_A^{1/2}}{\left[2^{1/2} \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/4} \right]}$$
$$= (2N_A)^{1/2} \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/4} e^{\left(\frac{E_V - E_A}{2kT}\right)}$$
(2.26)

From the above equation, it is seen that the density of holes in the valence band is proportional to the square root of the acceptor concentration N_A at moderately low temperatures.

VARIATION OF CARRIER CONCENTRATION WITH TEMPERATURE 2.12

Consider an N-type semiconductor doped with donor (Antimony, Sb) concentration of N_D donors per unit volume in which $N_D > n_i$. Figure 2.10 shows the thermal generation of electrons in an N-type semiconductor for different temperatures. At very low temperature, the conductivity of semiconductor is very low because the donor atoms are not ionized due to their insufficient thermal vibrational energy. As the temperature increases, some of the donors get ionized and donate their electrons to the conduction band, as shown in Fig. 2.10(a). Here, thermal excitation requires very high energy to break the covalent bonding of silicon atoms to move the electrons from valence band to conduction band. Therefore, thermal generation involves exciting electrons from E_D to E_C with the small ionization energy of donors i.e., $\Delta E = E_C - E_D (\ll E_G)$.

For *N*-type semiconductor, the electron density at low temperature is

$$n = \left(\frac{1}{2}N_C N_D\right)^{1/2} e^{\frac{-\Delta E}{2kT}}$$
(2.27)

For intrinsic semiconductor, the carrier concentration is

$$n_i = (N_C N_V)^{1/2} e^{\frac{-E_G}{2kT}}$$
(2.28)

Comparing Eqns. (2.27) and (2.28), it is seen that at low temperature, in an intrinsic semiconductor, thermal generation of carriers takes place across the bandgap E_G from valence band E_V to conduction band E_C whereas in an *N*-type semiconductor, the thermal excitation takes place across ΔE i.e., from E_D to E_C i.e., due to its doping with donor concentration, N_D .



Fig. 2.10 Excitation of electrons in N-type semiconductor with increase in temperature $(a)T < T_s(b)T_s > T < T_i$ and $(c)T > T_i$

Further increase in temperature leads to ionization of all the donors and the concentration of electrons becomes equal to the concentration of donor atoms, i.e., $n \approx N_D$ as shown in Fig. 2.10(b). At very high temperature, the thermal vibrational energy is very strong such that the covalent bonds of silicon atoms get broken and this results in the domination of thermal generation of carriers across E_G . As a result of breakage of covalent bonds, electrons move from valence band to conduction band due to thermal excitation, as shown in Fig. 2.10(c). This leads to the generation of an equal amount of hole concentration in the valence band. Hence, at very high temperature, the *N*-type semiconductor behaves as an intrinsic semiconductor i.e., the electron concentration of *N*-type will be equal to the intrinsic concentration n_i as given in Eqn. (2.28).

The variation in electron concentration with increase in temperature is distinguished using the following three regions:

- (i) Low temperature range: For $T < T_s$, more donor atoms get ionized. This ionization of donor atoms continues until the saturation temperature T_s is reached. When all donor atoms are ionized, there results in the saturation of ionized donor concentration. This temperature range is called the *ionization range*. The concentration of electrons in this range is given by Eqn. (2.27).
- (ii) Medium temperature range: For $T_s < T < T_i$, most of the donor atoms are ionized i.e., $n \approx N_D$. This condition remains the same till $T = T_i$, when the intrinsic concentration n_i becomes equal to the donor concentration N_D since n_i is temperature dependent. This temperature range is called the *extrinsic range*. The applications of *PN* junction devices make use of the *N*-type doping properties of the semiconductor in this range.

(iii) *High temperature range:* When the temperature increases further above the intrinsic range i.e., $T > T_i$, there is high thermal excitation energy which leads to the increase in the concentration of electrons. As a result, the electron concentration equals the intrinsic concentration i.e., $n = n_i$. This temperature range is called the *intrinsic range*. Further, the movement of electrons to the conduction band results in the creation of holes in the valence band i.e., p = n.

Figure 2.11 shows the temperature dependence of electron concentration n in an N-type semiconductor. Here, the graph is plotted with $\ln(n)$ for inverse variations $\ln(n)$

in temperature T^{-1} . The graph is almost a straight line with an ionization slope of $-\Delta E / 2k$ for very low temperature range. The temperature dependence is negligible in the intermediate range ($n \approx N_D$) whereas in the high temperature range, the graph has a very steep slope of $-E_G / 2k$.

Figure 2.12 shows the variation of intrinsic carrier concentration n_i for Ge, Si and GaAs with temperature. The graph is plotted with log (n_i) for inverse variations in temperature T^{-1} . Here, the slope of the lines is a measure of bandgap energy E_G . As the temperature increases, the bandgap energy also increases which gives rise to more electron-hole generation.



Fig. 2.11 Variation of electron concentration with temperature in N-type semiconductor



Fig. 2.12 Variation of intrinsic concentration with temperature

2.13 VARIATION OF FERMI LEVEL WITH TEMPERATURE AND IMPURITY CONCENTRATION

Figure 2.13(a) shows the variation of Fermi level E_F with temperature in *N*-type semiconductors. From Eqn. (2.19), it is seen that the Fermi level increases with increase in temperature. An increase of temperature causes more and more donor atoms to ionize. Further increase in temperature breaks the covalent bonds and this result in the generation of *electron-hole* pairs. So, the material tends to behave in an intrinsic manner. The Fermi level of *N*-type semiconductor gradually moves closer to the intrinsic Fermi level E_{Fi} . At higher temperatures, the semiconductor material begins to lose its extrinsic characteristics and acts like an intrinsic semiconductor. At lower temperatures, freeze-out occurs and the Fermi level goes above E_D for *N*-type semiconductor as shown in Fig. 2.13(a).

Figure 2.13(a) also shows the variation of Fermi level E_F with donor impurity concentration N_D in *N*-type semiconductors. As the donor concentration increases, the Fermi level moves closer towards the conduction band in *N*-type semiconductor. In case, the semiconductor is heavily doped, it will exhibit intrinsic behaviour only after all the donor atoms are ionized. $N_A = 10^{24}$ atoms/m³



concentration in (a) N-type and (b) P-type semiconductors

Figure 2.13(b) shows the variation of Fermi level E_F with temperature in *P*-type semiconductors. From Eqn. (2.24), it is seen that E_F decreases with increase in temperature. As the temperature is increased, more and more acceptor atoms are ionized as shown in Fig. 2.13(b). If the temperature is increased further, *electron-hole* pairs are generated due to breaking of covalent bonds. This increases the number of electrons as well. The number of electrons and holes become almost equal and the material behaves like an intrinsic semiconductor. Consequently, the Fermi level in a *P*-type semiconductor moves closer to the intrinsic Fermi level E_{Fi} . At higher temperatures, the semiconductor material begins to lose its extrinsic characteristics and acts like an intrinsic semiconductor. At lower temperatures, freeze-out occurs and the Fermi level goes below E_A for *P*-type semiconductor as shown in Fig. 2.13(b).

Figure 2.13(b) also shows the variation of Fermi level E_F with acceptor impurity concentration N_A in *P*-type semiconductors. As the acceptor concentration increases, the Fermi level moves closer towards the valence band in *P*-type semiconductor. In case, the semiconductor is heavily doped, it will exhibit intrinsic behaviour only after all the acceptor atoms are ionized.

2.14 VARIATION IN SEMICONDUCTOR PARAMETERS WITH TEMPERATURE

As the semiconductors operate on a wide range of temperature, the variation of semiconductor parameters such as intrinsic concentration (n_i) , mobility (μ) , conductivity (σ) and energy gap (E_G) with temperature are important and are discussed below.

Intrinsic concentration (n_i) We know that the electron concentration,

$$n = N_C e^{-(E_C - E_F)/kT}$$

and hole concentration,

$$p = N_V e^{-(E_F - E_V)/kT}$$

Therefore, the product of electron-hole concentrations,

$$np = N_C N_V e^{-(E_C - E_V)/kT}$$
$$= N_C N_V e^{-(E_G)/kT}$$

According to mass-action law, $np = n_i^2$. Hence, it is clear that the intrinsic concentration, n_i , is independent of the Fermi level but depends on temperature and energy gap, $E_G = E_C - E_V$. The above equation is valid for either an extrinsic or intrinsic material.

Also, in an intrinsic semiconductor, as temperature rises the intrinsic concentration (n_i) increases and the conductivity (σ) also increases. The intrinsic concentration n_i is very sensitive to temperature and is given by

$$n_i^2 = A_0 T^3 e^{-E_{GO}/kT}$$

where $A_o =$ a constant, independent of temperature

T =temperature (in K)

 E_{GO} = forbidden energy gap at K (in eV) and

k = Boltzmann constant (in eV/K).

The increase in n_i^2 with temperature has an effect on the charge densities in extrinsic semiconductor. In an *N*-type semiconductor, the number of free electrons (*n*) does not change appreciably with the increase in temperature, while the number of holes increases. In a *P*-type semiconductor, the number of free electrons (*n*) increases with the increase in temperature, while the number of holes (*p*) remains constant.

Mobility (μ) The mobility (μ) of an intrinsic semiconductor varies as T^{-m} over a temperature (T) range of 100 to 400 K. For silicon, m = 2.5 for electrons and 2.7 for holes. Similarly, for germanium m = 1.66 for electrons and 2.33 for holes. The mobility of an intrinsic semiconductor decreases with the rise in temperature. Also, the mobility is a function of electric field intensity (EV) and it remains constant only if $E < 10^3$ m in

N-type silicon. For $10^3 < E < 10^4$ V/m, μ_n varies approximately as $E^{-1/2}$. For higher fields, $\mu_n = \frac{V_d}{E}$ and hence,

 μ_n varies inversely as E and the carrier speed approaches a constant value of about 10⁵ m/s.

Conductivity (σ) The conductivity (σ) of an intrinsic semiconductor depends upon the number of holeelectron pairs and mobility. The number of hole-electron pairs increases with the rise in temperature, while its mobility decreases. However, the conductivity of an intrinsic semiconductor increases with the increase in temperature. The conductivity at any temperature (T K) is given by

$$\sigma = \sigma_o \left[1 + \alpha \left(T - T_o \right) \right]$$

where α is the temperature coefficient.

The conductivity of extrinsic semiconductor decreases with the rise in temperature, as the number of majority carriers is almost constant and mobility decreases.

Energy gap The energy gap decreases with the increase in temperature and is given by

$$E_G(T) = E_{Go} - \beta T$$

where $\beta = a$ constant, whose value depends upon the nature of the material. Its value for silicon is 3.6×10^{-4} and for germanium 2.23×10^{-4} , and E_{Go} = energy gap at 0 K. Its value for silicon is 1.21 eV and for germanium, 0.785 eV.

EXAMPLE 2.21

Calculate the position of Fermi-level E_F and σ at 300 K for Ge crystal containing 5×10^{22} arsenic atoms/m³. Also calculate σ if the mobility of the electrons is 0.39 m² / V·s. Assume $m_e^* = m_0$

Solution Given $n = 5 \times 10^{22} / \text{m}^3$ and $\mu_n = 0.39 \text{ m}^2 / \text{V} \cdot \text{s}$.

In N-type semiconductor, the density of electrons 'n' in the conduction band is

$$n = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} e^{\left(\frac{E_F - E_C}{kT}\right)}$$

i.e.,

$$e^{\left(\frac{E_F - E_C}{kT}\right)} = 2.5 \times 10^{-3}$$

Taking natural logarithm on both sides of the above equation, we get

$$\frac{E_F - E_C}{kT} = \ln(2.5 \times 10^{-3}) = -5.99$$

 $5 \times 10^{22} = 2.5 \times 10^{25} e^{\left(\frac{E_F - E_C}{kT}\right)}$

or

$$E_C - E_F = \frac{1.38 \times 10^{-23} \times 300}{1.602 \times 10^{-19}} \times 5.99 = 0.16 \text{ eV}$$

Conductivity,

$$\sigma = nq\mu_n = 5 \times 10^{22} \times 1.602 \times 10^{-19} \times 0.39 = 3.12 \times 10^3 (\Omega \cdot m)^{-10}$$

EXAMPLE 2.22

A sample of intrinsic germanium at room temperature has a carrier concentration of 2.4×10^{19} m³. It is doped with antimony at a rate of one antimony atom per million atoms of germanium. If the concentration of the germanium atoms is 4×10^{28} m³, determine the hole concentration.

Solution Given $ni = 2.4 \times 10^{19} \text{ m}^3$ and $N = 4 \times 10^{28} / \text{m}^3$. Density of donor atoms, $N_D = \frac{N}{10^6} = \frac{4 \times 10^{28}}{10^6} = 4 \times 10^{22} / \text{m}^3$ Free electron concentration, $n = N_D = 4 \times 10^{20} / \text{m}^3$ Hole concentration,

$p = \frac{n_i^2}{n} = \frac{(2.4 \times 10^{19})^2}{4 \times 10^{22}} = 1.4 \times 10^{16} / \text{m}^3$

EXAMPLE 2.23

A sample of intrinsic silicon at room temperature has a carrier concentration of 1.5×10^{16} / m³. A donor impurity is added to the extent of 1 donor atom per 10^8 atoms of silicon. If the concentration of silicon atoms is 5 × 10²⁸ atoms/m³, determine the resistivity of the material. (Given $\mu_n = 0.135 \text{ m}^2 / \text{V} \cdot \text{s.}$ and $\mu_n = 0.048 \text{ m}^2 / \text{V} \cdot \text{s}$).

Solution Given $n_i = 1.5 \times 10^{16} / \text{m}^3$ and $N = 5 \times 10^{28} / \text{m}^3$.

Density of donor atoms, $N_D = \frac{N}{10^8} = \frac{5 \times 10^{28}}{10^8} = 5 \times 10^{20} / \text{m}^3$

Free electron concentration,
$$n = N_D = 5 \times 10^{20} / \text{m}^3$$

p

Hole concentration.

$$=\frac{n_i^2}{n}=\frac{(1.5\times10^{10})^2}{5\times10^{20}}=4.5\times10^{11}\,/\,\mathrm{m}^3$$

1

Resistivity of the material, $\rho = \frac{1}{2}$

$$= \frac{1}{\sigma} = \frac{1}{q(n\mu_n + p\mu_p)}$$
$$= \frac{1}{(1.602 \times 10^{-19})(5 \times 10^{20} \times 0.135 + 4.5 \times 10^{11} \times 0.048)} = 0.092 \ \Omega \cdot m$$

EXAMPLE 2.24

Silicon has a conductivity of only $5 \times 10^{-4} (\Omega \cdot m)^{-1}$ in its pure form. The conductivity becomes 200 $(\Omega \cdot m)^{-1}$ when the pure silicon is doped with aluminium to produce a P-type semiconductor. Calculate the impurity concentration. Assume $\mu_p = 0.05 \text{ m}^2 / \text{V} \cdot \text{s}$.

Solution Given conductivity of impure silicon, $\sigma = 200 (\Omega \cdot m)^{-1}$.

We know that the conductivity, $\sigma = pq\mu_p = N_A q\mu_p$

The acceptor impurity concentration is

$$N_A = \frac{\sigma}{q\mu_p} = \frac{200}{1.602 \times 10^{-19} \times 0.05} = 2.5 \times 10^{22} \text{ atoms/m}^3$$

(2)

EXAMPLE 2.25

In a *P*-type germanium, $n_i = 2.1 \times 10^{19} / \text{m}^3$, density of boron 4.5×10^{23} atoms/m³, the electron and hole mobilities are 0.4 and 0.2 m² / V ·s respectively. What is the conductivity before and after the addition of boron atoms? (AU May 2003)

Solution Given $\mu_n = 0.4 \text{ m}^2 / \text{V} \cdot \text{s} \ \mu_n = 0.2 \text{ m}^2 / \text{V} \cdot \text{s} \ p = 4.5 \times 10^{23} \text{ atoms/m}^3$, and $n_i = 2.1 \times 10^{19} / \text{m}^3$.

Before the addition of boron atoms, germanium is a pure (intrinsic) semiconductor and the conductivity is due to both electrons and holes.

Hence,

$$\sigma_i = n_i q(\mu_n + \mu_p) = 2.1 \times 10^{19} \times 1.602 \times 10^{-19} (0.4 + 0.2) = 2.016 (\Omega \cdot m)^{-1}$$

After the addition of boron atoms, germanium becomes P-type (extrinsic) semiconductor and the conductivity is due to majority carriers i.e., holes only.

Therefore.

$$\sigma = pq\mu_p = 4.5 \times 10^{23} \times 1.602 \times 10^{-19} \times 0.2 = 1.44 \times 10^4 (\Omega \cdot m)^{-1}$$

EXAMPLE 2.26

In an N-type semiconductor, the concentration of electron is 2×10^{22} / m³. Its electrical conductivity is 112 $(\Omega \cdot m)^{-1}$. Calculate the mobility of electrons. (AU Nov 2005)

Solution Given $n = 2 \times 10^{22}$ atoms/m³ and $\sigma = 112 (\Omega \cdot m)^{-1}$.

For *N*-type semiconductor, $\sigma = n_{\alpha}\mu_{n}$

Hence, mobility of electrons is $\mu_n = \frac{\sigma}{na} = \frac{112}{2 \times 10^{22} \times 1.602 \times 10^{-19}} = 0.035 \,\text{m}^2 \,/\,\text{V} \cdot \text{s}$

EXAMPLE 2.27

In an N-type semiconductor, the Fermi level is 0.3 eV below the conduction level at a room temperature of 300 K. If the temperature is increased to 360 K, determine the new position of the Fermi level.

Solution The Fermi level in an *N*-type material is given by

Therefore,

 $E_F = E_C - kT \ln \frac{N_C}{N_D}$

At T = 300 K,

 $(E_C - E_F) = kT \ln \frac{N_C}{N_D}$ $0.3 = 300 k \ln \frac{N_C}{N_D}$ $E_C - E_{F1} = 360 k \ln \frac{N_C}{N_D}$ (1)

Similarly,

Dividing Eqn. (2) by Eqn. (1), we get

Therefore,

$$\frac{E_C - E_{F1}}{0.3} = \frac{360}{300}$$

$$E_C - E_{F1} = \frac{360}{300} \times 0.3 = 0.36 \text{ eV}$$

Hence, the new position of the Fermi level lies 0.36 eV below the conduction level.

In a *P*-type semiconductor, the Fermi level is 0.3 eV above the valance band at a room temperature of 300 K. Determine the new position of the Fermi level for temperatures of (a) 350 K, and (b) 400 K.

Solution The Fermi level in a *P*-type material is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A}$$

Therefore,

$$(E_F - E_V) = kT \ln \frac{N_V}{N_A}$$

At
$$T = 300$$
 K, $0.3 = 300 k \ln \frac{N_V}{N_A}$

(a) At T = 350 K, $(E_{F1} - E_V) = 350k \ln \frac{N_V}{N_A}$ Hence, from the above equation

$$\frac{E_{F1} - E_V}{0.3} = \frac{350}{300}$$

Therefore,

$$E_{F1} - E_V = \frac{350}{300} \times 0.3 = 0.35 \text{ eV}$$

(b) At T = 400 K, $(E_{F2} - E_V) = 400 k \ln \frac{1}{N_A}$ Hence, from the above equation,

$$\frac{E_{F2} - E_V}{0.3} = \frac{400}{300}$$
$$E_{F2} - E_V = \frac{400}{300} \times 0.3 = 0.4 \text{ eV}$$

Therefore,

EXAMPLE 2.29

In an *N*-type semiconductor, the Fermi level lies 0.2 eV below the conduction band. Find the new position of Fermi level if the concentration of donor atoms is increased by a factor to (a) 4 and (b) 8. Assume kT = 0.025 eV.

Solution In an N-type material, the concentration of donor atoms is given by

$$N_D = N_C e^{-(E_C - E_F)/kT}$$

Let initially $N_D = N_{DO}$, $E_F = E_{FO}$ and $E_C - E_{FO} = 0.2$ eV

Therefore,

(a) When $N_D = 4N_{DO}$ and $E_F = E_{F1}$, then

 $4N_{DO} = N_C e^{-(E_C - E_{Fl})/0.025} = N_C e^{-40 (E_C - E_{Fl})}$ $4 \times N_C e^{-8} = N_C e^{-40(E_C - E_{Fl})}$

 $N_{DQ} = N_C e^{-0.2/0.025} = N_C e^{-8}$

Therefore,

 $4 = e^{-40(E_C - E_{F_1}) + 8}$

Therefore,

Taking natural logarithm on both sides, we get
$$\ln 4 = -40 (E_C - E_{F1}) + 8$$

1.386 = -40(E_C - E_{F1}) + 8

 $E_C - E_{F1} = 0.165 \text{ eV}$ Therefore.

(b) When
$$N_D = 8N_{DO}$$
 and $E_F = E_{F2}$, then
 $\ln 8 = -40 (E_C - E_{F2}) + 8$
 $2.08 = -40 (E_C - E_{F2}) + 8$
Therefore, $E_C - E_{F2} = 0.148 \text{ eV}$

EXAMPLE 2.30

In a P-type semiconductor, the Fermi level lies 0.4 eV above the valence band. Determine the new position of Fermi level if the concentration of acceptor atoms is multiplied by a factor of (a) 0.5 and (b) 4. Assume kT = 0.025 eV.

Solution In a *P*-type material, the concentration of acceptor atoms is given by

$$N_A = N_V e^{-(E_F - E_V)/kT}$$

Let initially $N_A = N_{AO}$, $E_F = E_{FO}$ and $E_{FO} - E_V = 0.4$ eV

Therefore,

$$N_{AQ} = N_V e^{-0.4/0.025} = N_V e^{-16}$$

(a) When $N_A = 0.5$, N_{AO} and $E_F = E_{F1}$, then $0.5N_{AO} = N_V e^{-(E_{F1} - E_V)/0.025} = N_V e^{-40 (E_{F1} - E_V)}$

 $0.5 \times N_V e^{-16} = N_V e^{-40(E_{F1} - E_V)}$ $0.5 = e^{-40(E_{F1} + E_V) + 16}$

Therefore.

Taking natural logarithm on both sides, we get

 $\ln (0.5) = -40(E_{F1} - E_V) + 16$

Therefore,

$$E_{F1} - E_V = 0.417 \text{ eV}$$

(b) When $N_A = 4N_{AO}$ and $E_F = E_{F2}$, then

 $\ln 4 = -40(E_{F2} - E_V) + 16$

 $E_{F2} - E_V = 0.365 \text{ eV}$ Therefore,

EXAMPLE 2.31

Consider a silicon *PN* junction at T = 300 K so that $n_i = 1.5 \times 10^{10}$ cm⁻³. The *N* type doping is 1×10^{10} cm⁻³ and a forward bias of 0.6 V is applied to the PN junction. Calculate the minority hole concentration at the edge of the space charge region. (AU April/May 2015)

Solution Given T = 300K, $n_i = 1.5 \times 10^{10}$ cm⁻³, $n \times 1 \times 10^{10}$ cm⁻³ and $V_F = 0.6$ V.

We know that

 $n \cdot p = n_i^2$ $p = \frac{n_i^2}{n}$

Therefore, the concentration of holes,

$$=\frac{(1.5\times10^{10})^2}{1\times10^{10}}=2.25\times10^{10}\,\mathrm{cm}^{-3}$$

EXAMPLE 2.32

Find the conductivity of silicon (a) in intrinsic condition at a room temperature of 300 K, (b) with donor impurity of 1 in 10^8 , (c) with acceptor impurity of 1 in 5×10^7 and (d) with both the above impurities present simultaneously. Given that n_i for silicon at 300 K is 1.5×10^{10} cm⁻³, $\mu_n = 1300$ cm²/V-s, $\mu_p = 500$ cm²/V-s, number of Si atoms per cm³ = 5×10^{22} .

Solution

(a) In intrinsic condition, $n = p = n_i$

Hence,

 $\sigma_i = qn_i (\mu_n + \mu_p)$ = (1.602 × 10⁻¹⁹) (1.5 × 10¹⁰) (1300 + 500) = 4.32 × 10⁻⁶ S/cm

(b) Number of silicon atoms/cm³ = 5×10^{22}

Hence, Further,

Therefore,

 $N_D = \frac{5 \times 10^{22}}{10^8} = 5 \times 10^{14} \text{ cm}^{-3}$ $n \approx N_D$ $p = \frac{n_i^2}{n} \approx \frac{n_i^2}{N_D}$ $= \frac{(1.5 \times 10^{10})^2}{5 \times 10^{14}} = 0.46 \times 10^6 \text{ cm}^{-3}$

Thus, $p \ll n$. Hence, p may be neglected while calculating the conductivity.

 $\sigma = nq\mu_n = N_D q\mu_n$ = (5 × 10¹⁴) (1.602 × 10⁻¹⁹) (1300) = 0.104 S/cm $N_A = \frac{5 × 10^{22}}{5 × 10^7} = 10^{15} \text{ cm}^{-3}$ $p \approx N_A$ $n_i^2 = n_i^2$

(c)

Further,

Hence,

$$n = \frac{n_i^2}{p} \approx \frac{n_i^2}{N_A}$$
$$= \frac{(1.5 \times 10^{10})^2}{10^{15}} = 2.25 \times 10^5 \,\mathrm{cm}^{-3}$$

Thus, p >> n. Hence, n may be neglected while calculating the conductivity. Hence, $\sigma = pq\mu_n = N_A q\mu_n$

$$= (10^{15} \times 1.602 \times 10^{-19} \times 500)$$
$$= 0.08 \text{ S/cm}$$

(d) With both types of impurities present simultaneously, the net acceptor impurity density is,

Hence,

$$N'_{A} = N_{A} - N_{D} = 10^{15} - 5 \times 10^{14} = 5 \times 10^{14} \text{ cm}^{-3}$$

$$\sigma = N'_{A} q\mu_{p}$$

$$= (5 \times 10^{14}) (1.602 \times 10^{-19}) (500)$$

$$= 0.04 \text{ S/cm}$$

EXAMPLE 2.33

Determine the resistivity of germanium (a) in intrinsic condition at 300 K, (b) with donor impurity of 1 in 10⁷, (c) with acceptor impurity of 1 in 10⁸, (d) with both the above impurities simultaneously. Given that for germanium at room temperature $n_i = 2.5 \times 10^{13}$ /cm³, $\mu_n = 3800$ cm²/V-Vs, $\mu_p = 1800$ cm²/V-Vs and a number of Germanium atoms/cm³ = 4.4 × 10²².

Solution

(a)	$n = p = n_i = 2.5 \times 10^{13} \text{ cm}^{-3}$		
(b)	Therefore, conductivity,	$\sigma = qn_i(\mu_n + \mu_p)$	
		$= (1.602 \times 10^{-19})(2.5 \times 10^{13})(3800 + 1800)$	
		= 0.0224 S/cm	
	Hence, resistivity,	$\rho = \frac{1}{\sigma} = \frac{1}{0.0224} = 44.64 \ \Omega$ -cm	
		$N_D = \frac{4.4 \times 10^{22}}{10^7} = 4.4 \times 10^{15} \mathrm{cm}^{-3}$	
	Also,	$n = N_D$	
	Therefore,	$p = \frac{n_i^2}{n} = \frac{n_i^2}{N_D}$	
		$=\frac{(2.5 \times 10^{13})^2}{4.4 \times 10^{15}} = 1.42 \times 10^{11} \text{ holes/cm}^3$	
	Here, as $n >> p, p$ can be neglected.		
(c)	Therefore, conductivity,	$\sigma = nq\mu_n = N_D q\mu_n$	
		= (4.4×10^{15}) (1.602×10^{-19}) (3800) = 2.675 S/cm	
	Hence, resistivity,	$\rho = \frac{1}{\sigma} = \frac{1}{2.675} = 0.374 \Omega\text{-cm}$	
		$N_A = \frac{4.4 \times 10^{22}}{10^8} = 4.4 \times 10^{14} \mathrm{cm}^{-3}$	
	Also,	$p = N_A$	
	Therefore,	$n = \frac{n_i^2}{p} = \frac{n_i^2}{N_A} = \frac{(2.5 \times 10^{13})^2}{4.4 \times 10^{14}} = 1.42 \times 10^{12} \text{ electrons/cm}^3$	
	Here, as $p >> n, n$ may be neglected. Then		
	Conductivity,	$\sigma = pq\mu_p = N_A q\mu_p$	

 $= (4.4 \times 10^{14}) (1.602 \times 10^{-19}) (1800) = 0.1267$ S/cm

Hence, resistivity,

 $\rho = \frac{1}{\sigma} = \frac{1}{0.1267} = 7.89 \ \Omega$ -cm

(d) With both p and n type impurities present,

$$N_D = 4.4 \times 10^{15} \text{ cm}^{-3} \text{ and } N_A = 4.4 \times 10^{14} \text{ cm}^{-3}$$

Therefore, the net donor density N'_D is

$$N'_D = (N_D - N_A) = (4.4 \times 10^{15} - 4.4 \times 10^{14}) = 3.96 \times 10^{15} \text{ cm}^{-3}$$

Therefore, effective $n = N'_D = 3.96 \times 10^{15} \text{ cm}^{-3}$

$$p = \frac{n_i^2}{N'_D} = \frac{(2.5 \times 10^{13})^2}{3.96 \times 10^{15}} = 1.578 \times 10^{11} \,\mathrm{cm}^{-3}$$

Here, again $p\left(=\frac{n_i^2}{N_D'}\right)$ is very small compared with N_D' and may be neglected in calculating the effective

conductivity.

Therefore,

$$\sigma = N_D' q \mu_n$$

= (3.96 × 10¹⁵) (1.602 × 10⁻¹⁹) (3800) = 2.408 S/cm
$$\rho = \frac{1}{\sigma} = \frac{1}{2.408} = 0.415 \ \Omega\text{-cm}$$

Hence, resistivity

EXAMPLE 2.34

A sample of silicon at a given temperature T in intrinsic condition has a resistivity of $25 \times 10^4 \Omega$ -cm. The sample is now doped to the extent of 4×10^{10} donor atoms/cm³ and 10^{10} acceptor atoms/cm³. Find the total conduction current density if an electric field of 4 V/cm is applied across the sample. Given that $\mu_n = 1250 \text{ cm}^2/\text{V-s}$, $\mu_n = 475 \text{ cm}^2/\text{V-s}$ at the given temperature.

Solution

$$\sigma_i = qn_i(\mu_n + \mu_p) = \frac{1}{25 \times 10^4}$$

$$n_i = \frac{\sigma_i}{q(\mu_n + \mu_p)} = \frac{1}{(25 \times 10^4) (1.602 \times 10^{-19}) (1250 + 475)}$$

$$= 1.45 \times 10^{10} \text{ cm}^{-3}$$

$$N_D (= n) = (4 \times 10^{10} - 10^{10}) = 3 \times 10^{10} \text{ cm}^{-3}$$

$$n^2 = (1.45 \times 10^{10})^2$$

Therefore,

$$p = \frac{n_i^2}{N_D} = \frac{(1.45 \times 10^{10})^2}{3 \times 10^{10}} = 0.7 \times 10^{10} \,\mathrm{cm}^{-3}$$

Hence,

Hence,

$$\sigma = q(n\mu_n + p\mu_p)$$

= (1.602 × 10⁻¹⁹) (3 × 10¹⁰ × 1250 + 0.7 × 10¹⁰ × 475)
= 6.532 × 10⁻⁶ S/cm

Therefore, total conduction current density,

$$J = \sigma E = 6.532 \times 10^{-6} \times 4 = 26.128 \times 10^{-6} \text{ A/cm}^2$$

EXAMPLE 2.35

Find the concentration (densities) of holes and electrons in *N*-type silicon at 300 K, if the conductivity is 300 S/cm. Also find these values for *P*-type silicon. Given that for Silicon at 300 K, $n_i = 1.5 \times 10^{10}$ /cm³, $\mu_n = 1300$ cm²/V-s and $\mu_p = 500$ cm²/V-s.

Solution

(a) Concentration in *N*-type silicon

Concentration of electrons,

The conductivity of an *N*-type silicon is $\sigma = qn\mu_n$

$$n = \frac{\sigma}{q\mu_n}$$

= $\frac{300}{(1.602 \times 10^{-19}) (1300)} = 1.442 \times 10^{18} \text{ cm}^{-3}$
, $p = \frac{n_i^2}{n} = \frac{(1.5 \times 10^{10})^2}{1.442 \times 10^{18}} = 1.56 \times 10^2 \text{ cm}^{-3}$

Hence concentration of holes,

(b) Concentration in *P*-type silicon

The conductivity of a *P*-type silicon is $\sigma = qp\mu_p$

Hence, concentration of holes
$$p$$
:

$$= \frac{\overline{q\mu_p}}{300}$$

= $\frac{300}{(1.602 \times 10^{-19}) (500)}$ = 3.75 × 10¹⁸ cm⁻³

and concentration of electrons,
$$n = \frac{n_i^2}{p} = \frac{(1.5 \times 10^{10})^2}{3.75 \times 10^{18}} = 0.6 \times 10^2 \text{ cm}^{-3}$$

EXAMPLE 2.36

A specimen of pure germanium at 300 K has a density of charge carriers 2.5×10^{19} /m³. It is doped with donor impurity atoms at the rate of one impurity atom every 106 atoms of germanium. All impurity atoms are supposed to be ionised. The density of germanium atom is 4.2×10^{28} atoms/m³. Calculate the resistivity of the doped germanium if electron mobility is 0.38 m^2 /V-s.

If the Germanium bar is 5×10^{-3} m long and has a cross sectional area of $(5 \times 10^{-6})^2$ m², determine its resistance and the voltage drop across the semiconductor bar for a current of 1 µA flowing through it.

Solution Density of added impurity atoms is

$$N_D = \frac{4.2 \times 10^{28}}{10^6} = 4.2 \times 10^{22} \text{ atoms/m}^3$$

 $n \approx N_D$

Also,

Therefore,

$$p = \frac{n_i^2}{n} = \frac{n_i^2}{N_D} = \frac{(2.5 \times 10^{19})^2}{4.2 \times 10^{22}} = 1.488 \times 10^6 \,\mathrm{m}^{-3}$$

Here, as $p \ll n$, p may be neglected.

Therefore,
$$\sigma = q N_D \mu_n$$

	$= (1.602 \times 10^{-19}) (4.2 \times 10^{22}) (0.38) = 2.554 \times 10^3 \text{ S/m}$
Therefore, resistivity,	$\rho = \frac{1}{\sigma} = \frac{1}{2.554 \times 10^3} = 0.392 \times 10^{-3} \Omega\text{-m}$
Resistance,	$R = \frac{\rho L}{A} = \frac{0.392 \times 10^{-3} \times 5 \times 10^{-3}}{(5 \times 10^{-6})^2} = 78.4 \text{ k}\Omega$
Voltage drop,	$V = RI = 78.4 \times 10^3 \times 10^{-6} = 78.4 \text{ mV}$

EXAMPLE 2.37

The resistivities of the two sides of an abrupt germanium diode are 2Ω cm (*P* side) and 1Ω cm (*N* side) at 300 K. The mobility of electrons and holes in germanium are $\mu_n = 3800$ cm²/V sec and $\mu_p = 1800$ cm²/V sec respectively. Calculate the height E_0 of the potential-energy barrier. (AU May/June 2012)

Solution Since the conductivities of *P* and *N* are given by $\sigma_{p0} = q\mu_p p_{p0}$ and $\sigma_{n0} = q\mu_n n_{n0}$, the concentration of holes and electrons in the *P*- and *N*-type germanium junction is given by

$$P_{p0} = \frac{\sigma_{p0}}{q\mu_p} = \frac{\frac{1}{2}\Omega \text{ cm}}{(1.602 \times 10^{-19})(1800 \text{ cm}^2/\text{V sec})}$$
$$= 1.74 \times 10^{15} \text{ cm}^{-3}$$
$$n_{n0} = \frac{\sigma_{n0}}{q\mu_n} = \frac{\frac{1}{2}\Omega \text{ cm}}{(1.602 \times 10^{-19})(3800 \text{ cm}^2/\text{V sec})}$$

$$= 1.64 \times 10^{15} \text{ cm}^{-3}$$

The concentration of hole in N-side of the function is given by

$$P_{n0} = \frac{n_i^2}{n_{n0}} = \frac{6.25 \times 10^{26}}{1.64 \times 10^{15}} = 3.8 \times 10^{11} \text{ cm}^{-3}$$

We know that, the contact potential is

$$V_0 = \frac{E_0}{q} = \frac{kT}{q} \ln \frac{p_{p0}}{p_{n0}}$$

Therefore,

$$E_0 = 0.0259 \times \ln\left(\frac{1.74 \times 10^{15}}{3.8 \times 10^{11}}\right) = 0.218 \text{ eV}$$
$$V_0 = \frac{E_0}{q} = \frac{0.218 \text{ eV}}{q} = 0.218 \text{ V}$$

2.15 CARRIER TRANSPORT: VELOCITY-ELECTRIC FIELD RELATIONS

When a semiconductor is subjected to an electric field, carrier transport occurs as a result of the acceleration experienced by the free charge carriers in the presence of the field. Electrons move in a direction opposite to the applied field whereas holes move in the direction of the applied field. However, both carriers undergo scattering from lattice deformations due to presence of impurities or lattice vibrations (phonons). This, in effect, reduces their momentum along the field direction. At steady state, these two opposing effects balance each other and each type of charge carrier attains a steady average velocity called its *drift velocity*. The response of the charge carriers to the applied field can be represented by a velocity-field relation. This relation is essentially nonlinear. The brief qualitative discussion of the velocity-field relation at low and high electric fields is given below.

2.15.1 Low field response: Mobility, Conductivity

At low electric fields, the relationship between velocity, mobility and conductivity of the charges, to their scattering rate can be understood using an approach based on the following assumptions:

- (i) Electrons (and holes) experience repeated collisions from various scattering sources (lattice imperfections, impurities etc). The average time interval between successive collisions is called the scattering time, τ .
- (ii) Electrons do not interact with each other.
- (iii) The motion of electrons in between collisions can be represented by the standard equation of motion with the assumption that immediately after collision, the average drift velocity of the electrons becomes zero i.e., they lose their excess energy and relax back to random thermal velocity distribution.

The resulting equation of motion can be solved (refer to section 1.4 of Chapter 1) to obtain the expression for drift velocity v_d acquired by the electrons of effective mass m^* when subjected to electric field of strength E which is as follows:

$$v_d = -\frac{e\tau E}{m*} \tag{2.29}$$

The negative sign in the above equation indicates that v_d for electrons is in opposite direction to applied field *E*. This relationship is the *velocity-field relation for low fields*.

From the definition of mobility,

$$v_d = \mu E$$

Therefore, the mobility is

$$\mu = \frac{v_d}{E}$$

The current density due to electron transport is

$$J = nev_d = \frac{ne^2 \tau E}{m^*}$$
(2.30)

where *n* is the density of electrons or the number of electrons per m^3 of the semiconductor and *e* is the charge of the electron.

From Ohm's law, $J = \sigma E$

Therefore, conductivity due to electrons is

$$\sigma = \frac{ne^2\tau}{m^*} = ne\mu_n \tag{2.31}$$

The conductivity due to holes can also be obtained in the same way. Therefore, in a semiconductor, with holes and electrons contributing to transport, the total conductivity can be given by the expression as

$$\sigma = ne\mu_n + pe\mu_p \tag{2.32}$$

The scattering time of doped semiconductors decreases as number of ionised dopants increases. It is inversely proportional to the number of ionised dopants, and hence, mobility of charge carriers in doped semiconductors is generally lower. In addition, mobility is also found to be directly proportional to $T^{-3/2}$ i.e., $\mu_n \sim T^{-3/2}$. This temperature dependence is unique for ionised impurity scattering dominant in doped semiconductors. As seen earlier, carrier density *n* has a term proportional to $T^{3/2}$. Therefore, in the expression for conductivity, which is a product of *n* and μ_n these dependencies cancel out and its only temperature dependence arises from the exponential term $e^{-\Delta E/2kT}$ in number density *n*.

Velocity-field curve in semiconductors

Figure 2.14 illustrates this behaviour for direct and indirect bandgap semiconductors. Velocity-field relation for low field shows that the drift velocity acquired by an electron increases linearly with increase in applied field i.e., it is directly proportional to the field. This low field region is shown as region I (denoted by b) in Fig. 2.14. The drift velocity is also inversely proportional to the effective mass of the carriers. In the case of holes, the effective mass will be higher and therefore, the velocities can be expected to be lower.



Fig. 2.14 Velocity-field curve in semiconductors

2.15.2 High Field Transport: Velocity-Field Relation

At strong electric fields at constant temperature, electrons acquire higher average energy and hence, higher drift velocity. But as drift velocity increases, electrons excite out-of phase vibrations in neighbouring atoms of the lattice (optical phonons) further increasing the scattering rate or decreasing mean time between successive collisions. This competitive effect lowers drift velocity and mobility, thereby, leading to saturation of drift velocity at high fields (depicted as region III in Fig. 2.14). Thus, in all semiconductors, the drift velocity increases linearly with field at low field strengths, but at sufficiently high field strengths, it eventually saturates and becomes independent of applied field strength. The energy is transferred to high energy phonons (optical phonons as they can be excited by IR radiations also). This has applications in generating microwaves.

In the case of direct bandgap materials, as carriers gain enough energy comparable to inter-valley separation in conduction band, they get excited to the high effective mass upper valley from the low effective mass lower energy state (refer to band bending and effective masses in the energy band diagrams). High effective mass implies that drift velocities drop and this results in a negative resistance region in the velocity-field curve for direct bandgap materials (depicted as region II in Fig. 2.14).

At very high field strengths (~100 kV/cm), the high energy gained by the electrons exceeds bandgap energy, and is dissipated through impact ionization generating additional electrons through avalanche breakdown and therefore, average drift velocity remains constant (depicted as region IV in Fig. 2.14). This process is seen in zener diodes. If the device dimension is smaller than the scattering time of the electrons/holes, then an increase in field strength will result in higher drift velocities than those seen in larger devices. This phenomenon of overshoot of drift velocity generally observed in direct bandgap materials with low effective mass carriers leads to higher conductivity. This high conductivity that can be seen in smaller devices (nano devices) is called ballistic conductivity and the transport phenomenon is called *ballistic transport*.

2.16 DRIFT AND DIFFUSION TRANSPORT

The determination of the concentration of the charged particles i.e., electrons and holes is important in understanding the electrical properties of a semiconductor material. The flow of current is mainly due to the movement of electrons and holes in a semiconductor. The process of movement of these charged particles is called *carrier transport*. The two basic carrier transport mechanisms in a semiconductor material are drift transport and diffusion transport.

The movement of charge due to electric fields in a semiconductor is called *drift transport* and the carrier transport through drift results in *drift current*. The flow of charge due to density gradients in a semiconductor is called *diffusion transport* and the carrier transport by diffusion results in *diffusion current*.

2.16.1 Drift Current

When an electric field is applied across the semiconductor material, the charge carriers attain a certain drift velocity v_d , which is equal to the product of the mobility of the charge carriers and the applied electric field intensity, *E*. The holes move towards the negative terminal of the battery and electrons move towards the positive terminal. This combined effect of movement of the charge carriers constitutes a current known as the *drift current*. Thus, the drift current is defined as the flow of electric current due to the motion of the charge carriers under the influence of an external electric field.

2.16.2 Drift Current Density

Figure 2.15 shows that a conducting wire of length l cm contains N electrons. If an electron travels a distance of l cm in the conductor in time T sec, the total number of electrons passing through any cross

section of wire per second is $\frac{N}{T}$.

The total current flowing through the wire with area of cross-section *A* is given by

$$I = \frac{Nq}{T} = \frac{Nqv_d}{l}$$



Fig. 2.15 Conducting wire to determine drift current density

where v_d is the drift velocity and time $T = \frac{l}{v_d}$. The drift current density J is defined as the current per unit

area of the conducting medium.

i.e.,
$$J = \frac{I}{A} = \frac{Nqv_d}{lA} \text{ A/cm}^2$$

Here, *lA* is the volume containing N electrons and the electron concentration, $n = \frac{N}{lA}$.

Therefore,
$$J = nqv_d = \rho v_d$$

where $\rho = nq$ is the charge density, in coulombs per cubic centimeter. Generally, the above current density can also be written as

$$J = nqv_d = nq\mu E = \sigma E$$

where *E* is the applied electric field intensity in V/cm, μ is the mobility of electrons in cm²/V-s, the drift velocity, $v_d = \mu E$ in cm/s, and the conductivity, $\sigma = nq\mu$ in S · m⁻¹. The equation $J = \sigma E$ is also called Ohm's law.

The drift current density due to the charge carriers such as free electrons and holes are the current passing through a square centimeter perpendicular to the direction of flow. The equation for the drift current density, J_n , due to free electrons is given by

$$J_n = qn\mu_n E \text{ A/cm}^2$$

and the drift current density, J_p , due to holes is given by

$$J_p = qp\mu_p E \text{ A/cm}^2$$

where n = number of free electrons per cubic centimetre

- p = number of holes per cubic centimetre
- μ_n = mobility of electrons in cm²/V-s
- μ_p = mobility of holes in cm²/V-s
- E = applied electric field intensity in V/cm
- q = charge of an electron = 1.602×10^{-19} coloumb.

2.16.3 Diffusion Current

It is possible for an electric current to flow in a semiconductor even in the absence of the applied voltage provided a concentration gradient exists in the material. A concentration gradient exists if the number of either electrons or holes is greater in one region of a semiconductor as compared to the rest of the region. In a semiconductor material, the charge carriers have the tendency to move from the region of higher concentration to that of lower concentration of the same type of charge carriers. Thus, the movement of charge carriers takes place resulting in a current called *diffusion current*. The diffusion current depends on the material of the semiconductor, type of charge carriers and the concentration gradient.

As indicated in Fig. 2.16(a), the hole concentration p(x) in a semiconductor bar varies from a high value to a low value along the *x*-axis and is constant in the *y*- and *z*-directions.



Fig. 2.16 (a) Excess hole concentration varying along the axis in an N-type semiconductor bar, (b) The resulting diffusion current

2.16.4 Diffusion Current Density

The diffusion current density J_p due to holes is given by

$$J_p = -qD_p \frac{dp}{dx} A/cm^2$$
(2.33a)

Since the hole density p(x) decreases with increasing x as shown in Fig. 2.16(b), dp/dx is negative and the minus sign in the above equation is needed in order that J_p has a positive sign in the positive x-direction.

The diffusion current density due to the free electrons, J_n , is given by

$$J_n = qD_n \frac{dn}{dx} A/cm^2$$
(2.33b)

where dn/dx and dp/dx are the concentration gradients for electrons and holes respectively, in the x-direction and D_n and D_p are the diffusion coefficients expressed in cm²/s for electrons and holes, respectively.

Total current The total current in a semiconductor is the sum of drift current and diffusion current. Therefore, for a *P*-type semiconductor, the total current per unit area, i.e., the total current density is given by

$$J_p = qp\mu_p E - qD_p \frac{dp}{dx}$$

Similarly, the total current density for an N-type semiconductor is given by

$$J_n = qn \ \mu_n \ E + qD_n \frac{dn}{dx}$$

2.17 EINSTEIN RELATIONSHIP FOR SEMICONDUCTOR

There exists a definite relationship between the mobility and diffusion coefficient of a particular type of charge carrier in the same semiconductor. The higher the value of mobility of a charge carrier, the greater will be its tendency to diffuse. The equation which relates the mobility μ and the diffusion coefficient *D* is known as the *Einstein Relationship*. The Einstein relationship is expressed as

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = \frac{kT}{q} = V_T \tag{2.34}$$

The importance of Einstein relationship is that it can be used to determine D_p (or D_n), if the mobility of holes (or electrons) is measured experimentally. For an intrinsic silicon, $D_p = 13 \text{ cm}^2/\text{s}$ and $D_n = 34 \text{ cm}^2/\text{s}$. For an intrinsic germanium, $D_p = 47 \text{ cm}^2/\text{s}$ and $D_n = 99 \text{ cm}^2/\text{s}$.

Diffusion length (L) As shown in Fig. 2.16, the excess hole or electron densities fall off exponentially with distance as a result of the recombination of these excess minority carriers with the majority carriers of the semiconductor. Here, the excess charge carriers have a finite life time, τ , before they are totally destroyed by recombination. The average distance that on excess charge carrier can diffuse during its life time is called the diffusion length *L*, which is given by

$$L = \sqrt{D\tau}$$

where *D* is the diffusion coefficient that may be related to the drift mobility, μ , through the Einstein relation as

$$D = \mu \frac{kT}{q}$$

If the transverse length of the semiconductor is greater than the diffusion length L, then the terminal currents are the recombination currents arising out of the recombination, as every electron lost by recombination is supplanted by the terminal electrode to maintain the charge neutrality.

EXAMPLE 2.38

Assume that in an *N*-type gallium arsenide semiconductor at T = 300 K, the electron concentration varies linearly from 1×10^{18} to 7×10^{17} cm⁻³ over a distance of 0.10 cm. Calculate the diffusion current density if the electron diffusion coefficient is $D_n = 225$ cm²/s.

Solution

Diffusion current density,
$$J_n = qD_n \frac{dn}{dx} = qD_n \frac{\Delta n}{\Delta x}$$

= 1.602×10⁻¹⁹×225× $\left(\frac{1 \times 10^{18} - 7 \times 10^{17}}{0.10}\right) = 108 \text{ A/cm}^2$

EXAMPLE 2.39

Find the diffusion coefficient of electron in silicon at 300 K if μ_n is 0.19 m² / V·s.

Solution Given $\mu_n = 0.19 \text{ m}^2 / \text{V} \cdot \text{s}$ at T = 300 K.

From Einstein's relation, $\frac{D_n}{\mu_n} = \frac{kT}{q}$, where charge of an electron, $q = 1.602 \times 10^{-19}$ C

Therefore, diffusion coefficient is

$$D_n = \frac{kT}{q} \mu_n = \frac{1.38 \times 10^{-23} \times 300}{1.602 \times 10^{-19}} \times 0.19 = 0.0049 \text{ m}^2 \text{ / s}$$

2.18 FORWARD AND REVERSE BIAS CHARACTERISTICS OF PN JUNCTION DIODE

2.18.1 PN Junction Diode in Equilibrium with no Applied Voltage

In a piece of semiconductor material, if one half is doped by P-type impurity and the other half is doped by N-type impurity, a PN junction is formed. The plane dividing the two halves or zones is called PNjunction. Here, the N-type material has high concentration of free electrons, while P-type material has high concentration of holes. Therefore, at the junction there is a tendency for the free electrons to diffuse over to the P-side and holes to the N-side. This process is called *diffusion*. As the free electrons move across the junction from N-type to P-type, the donor ions become positively charged. Hence a positive charge is built on the N-side of the junction. The free electrons that cross the junction uncover the negative acceptor ions by filling in the holes. Therefore, a net negative charge is established on the P-side of the junction. This net negative charge on the P-side prevents further diffusion of electrons into the P-side. Similarly, the net positive charge on the N-side repels the hole crossing from P-side to N-side. Thus a barrier is set-up near the junction which prevents further movement of charge carriers, i.e., electrons and holes. As a consequence of the induced electric field across the depletion layer, an electrostatic potential difference is established between P- and *N*-regions, which is called the potential barrier, junction barrier, diffusion potential, or contact potential, V_o . The magnitude of the contact potential V_o varies with doping levels and temperature. V_o is 0.3 V for germanium and 0.72 V for silicon.

The electrostatic field across the junction caused by the positively charged *N*-type region tends to drive the holes away from the junction and negatively charged *P*-type region tends to drive the electrons away from the junction. The majority holes diffusing out of the *P*-region leave behind negatively charged acceptor atoms bound to the lattice, thus exposing negative space charge in a previously neutral region. Similarly, electrons diffusing from the *N*-region expose positively ionised donor atoms, and a double space charge layer builds up at the junction as shown in Figs. 2.17(a) and (c).

It is noticed that the space-charge layers are of opposite sign to the majority carriers diffusing into them, which tends to reduce the diffusion rate. Thus, the double space of the layer causes an electric field to be set up across the junction directed from *N*- to *P*-regions, which is in such a direction to inhibit diffusion of majority electrons and holes, as illustrated in Fig. 2.17(a) and (d). The shape of the charge density, ρ , depends upon how the diode is doped, Thus, the junction region is depleted of mobile charge carriers. Hence, it is called the depletion region (layer), the space charge region, or the transition region. The depletion region is of order 0.5 µm thick. There are no mobile carriers in this very narrow depletion layer. Hence no current flows across the junction and the system is in equilibrium. To the left of this depletion layer, the carrier concentration is $p \approx N_A$, and to its right it is $n \approx N_D$.

Calculation of depletion width Let us now consider the width of the depletion region in the junction of Fig. 2.17. The region contains space charge due to the fact that, donors on the *N*-side and acceptors on the *P*-side have lost their accompanying electrons and holes. Hence, an electric field is established which, in turn, causes a difference in potential energy, qV_o , between the two parts of the specimen. Thus, a potential is built up across the junction and Fig. 2.17(e) represents the variation in potential. Here, *P*-side of the junction is at a lower potential than the *N*-side which means that the electrons on the *P*-side have a great potential energy.

In this analysis, let us consider an *alloy junction* in which there is an abrupt change from acceptor ions on *P*-side to donor ions on *N*-side. Assume that the concentration of electrons and holes in the depletion region is negligible and that all of the donors and acceptors are ionised. Hence, the regions of space charge may be described as

$$\rho = \begin{cases} -qN_{A}, & 0 > x > X_{1} \\ -qN_{D}, & X_{2} > x > 0 \\ 0, & elsewhere \end{cases}$$

where ρ is the space charge density, as indicated in Fig. 2.17(c)(i). The axes have been chosen in Fig. 2.17(e) in such a way that V_1 and X_1 have negative values. The potential variation in the space charge region can be calculated by using Poisson's equation, which is given by

$$\nabla^2 V = -\frac{\rho(x, y, z)}{\varepsilon_0 \varepsilon_r}$$

where \mathcal{E}_r is the relative permittivity. The relevant equation for the required one-dimensional problem is

$$\frac{d^2 V}{dx^2} = -\frac{\rho}{\varepsilon_0 \varepsilon_r}$$



Fig. 2.17 Formation of PN junction

Applying the above equation to the *P*-side of the junction, we get

$$\frac{d^2 V}{dx^2} = \frac{qN_A}{\varepsilon_0 \varepsilon_r}$$

Integrating twice, we get

As $V = V_1$ at $x = X_1$, we have

$$V = \frac{qN_A x^2}{2\varepsilon_0 \varepsilon_r} + Cx + D$$

where C and D are the constants of integration.

From the Fig. 2.17(e), we have V = 0 at x = 0, and hence D = 0. When $x < X_1$ on the *P*-side, the potential is constant, so that $\frac{dV}{dx} = 0$ at $x = X_1$. Hence,

$$C = -\frac{qN_A}{\varepsilon_o \varepsilon_r} \cdot X_1$$
$$V = \frac{qN_A x^2}{2\varepsilon_o \varepsilon_r} - \frac{qN_A}{\varepsilon_o \varepsilon_r} \cdot X_1 \cdot x$$

Therefore,

 $V = \frac{qN_A}{\varepsilon_o \varepsilon_r} \left(\frac{x^2}{2} - X_1 \cdot x \right)$

i.e.

$$V_1 = -\frac{qN_A}{2\varepsilon_o \,\varepsilon_r} \cdot X_1^2$$

If we apply the same procedure to the N-side, we get

$$V_2 = \frac{qN_D}{2\varepsilon_o \varepsilon_r} \cdot X_2^2$$

Therefore, the total built-in potential or the contact potential is V_o , where

$$V_{o} = V_{2} - V_{1} = \frac{q}{2\varepsilon_{o} \varepsilon_{r}} \left(N_{A} X_{1}^{2} + N_{D} X_{2}^{2} \right)$$

We know the fact that the positive charge on the N-side must be equal in magnitude to the negative charge on the P-side for the neutral specimen. Hence,

$$N_A X_1 = -N_D X_2$$

and substituting this relationship in the above equation and using the fact that X_1 is a negative quantity, we get

$$X_{1} = -\left[\frac{2\varepsilon_{o} \varepsilon_{r} V_{o}}{qN_{A} \left(1 + \frac{N_{A}}{N_{D}}\right)}\right]^{1/2}$$

Similarly,

$$X_{2} = \left[\frac{2\varepsilon_{o} \varepsilon_{r} V_{o}}{qN_{D} \left(1 + \frac{N_{D}}{N_{A}}\right)}\right]^{1/2}$$

The total depletion width, $W = X_2 - X_1$ and hence, $W^2 = X_1^2 + X_2^2 - 2X_1 X_2$, and then substituting for X_1 and X_2 from the above equations, we find

$$W = \left[\frac{2 \varepsilon_o \varepsilon_r V_o}{q} \left(\frac{N_A + N_D}{N_A N_D}\right)\right]^{1/2}$$

Here, in an *alloy junction*, the depletion width W is proportional to $(V_o)^{1/2}$.

In a *Grown Junction*, the charge density (ρ) varies linearly with distance (x) as shown in Fig. 2.17(c)(ii). If a similar analysis is carried for this junction, it is found that W varies as $(V_{\rho})^{1/3}$ instead of $(V_{\rho})^{1/2}$.

2.18.2 Under Forward Bias Condition

When positive terminal of the battery is connected to the *P*-type and negative terminal to the *N*-type of the *PN* junction diode, the bias applied is known as forward bias.

Operation As shown in Fig. 2.18, the applied potential with external battery acts in opposition to the

internal potential barrier and disturbs the equilibrium. As soon as equilibrium is disturbed by the application of an external voltage, the Fermi level is no longer continuous across the junction. Under the forward bias condition, the applied positive potential repels the holes in *P*-type region so that the holes move towards the junction and the applied negative potential repels the electrons in the *N*-type region and the electrons move towards the junction. Eventually, when the applied potential is more than the internal barrier potential, the depletion region and internal potential barrier disappear.

V–ICharacteristics of a diode under forward bias Under forward bias condition, the *V–I* characteristics of a *PN* unction diode are shown in Fig. 2.19. As the forward voltage (V_F) is increased, for $V_F < V_O$, the

forward current I_F is almost zero (region *OA*) because the potential barrier prevents the holes from *P*-region and electrons from *N*-region to flow across the depletion region in the opposite direction.



A feature worth to be noted in the forward characteristics shown in Fig. 2.19 is the cut in or threshold voltage (V_r) below which the current is very small. It is 0.3 V and 0.7 V for germanium and silicon,

respectively. At the cut in voltage, the potential barrier is overcome and the current through the junction starts to increase rapidly.

2.18.3 Under Reverse Bias Condition

When the negative terminal of the battery is connected to the *P*-type and positive terminal of the battery is connected to the *N*-type of the *PN* junction, the bias applied is known as reverse bias.



Fig. 2.18 PN junction under forward bias



Fig. 2.19 V-I characteristics of a diode under forward bias condition

Operation Under applied reverse bias as shown in Fig. 2.20, holes which form the majority carriers of the

P-side move towards the negative terminal of the battery and electrons which form the majority carrier of the *N*-side are attracted towards the positive terminal of the battery. Hence, the width of the depletion region which is depleted of mobile charge carriers increases. Thus, the electric field produced by applied reverse bias, is in the same direction as the electric field of the potential barrier. Hence, the resultant potential barrier is increased which prevents the flow of majority carriers in both directions; the depletion width, W, is proportional to $\sqrt{V_o}$ under reverse bias. Therefore, theoretically no current should flow in the external circuit. But in practice, a very small current of the order of a few microamperes flows under reverse bias as shown in Fig. 2.21. Electrons forming covalent bonds of the semiconductor atoms in the *P*- and *N*-type regions may absorb sufficient energy from heat and light to cause breaking of some covalent bonds. Hence electron-hole pairs are continually produced in both the regions. Under the reverse bias condition, the thermally generated holes in the *P*-region are attracted towards the negative terminal of the battery and the electrons in the N-region are attracted towards the positive terminal of the battery. Consequently, the minority carriers, electrons in the P-region and holes in







Fig. 2.21 V-I characteristics under reverse bias

the *N*-region, wander over to the junction and flow towards their majority carrier side giving rise to a small reverse current. This current is known as *reverse saturation current*, I_o . The magnitude of reverse saturation current mainly depends upon junction temperature because the major source of minority carriers is thermally broken covalent bonds.

For large applied reverse bias, the free electrons from the *N*-type moving towards the positive terminal of the battery acquire sufficient energy to move with high velocity to dislodge valence electrons from semiconductor atoms in the crystal. These newly liberated electrons, in turn, acquire sufficient energy to dislodge other parent electrons. Thus, a large number of free electrons are formed which is commonly called as an avalanche of free electrons. This leads to the breakdown of the junction leading to very large reverse current. The reverse voltage at which the junction breakdown occurs is known as *Breakdown Voltage*, V_{RD} .

2.18.4 PN Junction as a Diode

Figure 2.22 shows the current-voltage characteristics of *PN* junction. The characteristics of the *PN* junction vary enormously depending upon the polarity of the applied voltage. For a forward-bias voltage, the current increases exponentially with the increase of voltage. A small change in the forward-bias voltage increases the corresponding forward-bias current by orders of magnitude and hence the forward-bias *PN* junction will have a very small resistance. The level of current flowing across a forward-biased *PN* junction largely depends upon the junction area. In the reverse-bias direction, the current remains small, i.e., almost zero, irrespective of the magnitude of the applied voltage and hence the reverse-bias *PN* junction will have a high resistance. The reverse bias current depends on the area, temperature and type of semiconductor material.



Fig. 2.22 Ideal I-V characteristics of a PN junction diode

The semiconductor device that displays these *I-V* characteristics is called a *PN* junction diode. Figure 2.23 shows the *PN* junction diode with forward-bias and reverse-bias and their circuit symbols. The metal contacts are indicated with which the homogeneous *P*-type and *N*-type materials are provided. Thus two metal-semiconductor junctions, one at each end of the diode, are introduced. The contact potential across these junctions is approximately independent of the direction and magnitude of the current. A contact of this type is called an *ohmic contact*, which has low resistance. In the forward-bias, a relatively large current is produced by a fairly small applied voltage. In the reverse-bias, only a very small current, ranging from nanoamps to microamps is produced. The diode can be used as a voltage controlled switch, i.e., OFF for a reverse-bias voltage and ON for a forward-bias voltage.



(b) Reverse-biased PN junction diode and its circuit symbol

When a diode is reverse-biased by atleast 0.1V, the diode current is $I_R = -I_o$. As the current is in the reverse direction and is a constant, it is called the diode *reverse saturation current*. Real diodes exhibit reverse-bias current that are considerably larger than I_o . This additional current is called a *generation current* which is due to electrons and holes being generated within the space-charge region. A typical value of I_o may be 10^{-14} A and a typical value of reverse-bias current may be 10^{-9} A.

2.18.5 PN Junction Diode as Rectifier

A *PN* junction diode is a two terminal device that is polarity sensitive. When the diode is forward biased, the diode conducts and allows current to flow through it without any resistance, i.e., the diode is ON. When the diode is reverse biased, the diode does not conduct and no current flows through it, i.e., the diode is OFF, or providing a blocking function. Thus an ideal diode acts as a switch, either open or closed, depending upon the polarity of the voltage placed across it. The ideal diode has zero resistance under forward bias and infinite resistance under reverse bias.

2.18.6 Diode Ratings or Limiting Values of PN Junction Diode

The *PN* junction diode will perform satisfactorily only if it is operated within certain limiting values. They are the following:

- (a) Maximum Forward Current It is the highest instantaneous current under forward bias condition that can flow through the junction.
- (b) Peak Inverse Voltage (PIV) It is the maximum reverse voltage that can be applied to the *PN* junction. If the voltage across the junction exceeds PIV, under reverse bias condition, the junction gets damaged.
- (c) Maximum Power Rating It is the maximum power that can be dissipated at the junction without damaging the junction. Power dissipation is the product of voltage across the junction and current through the junction.
- (d) Maximum Average Forward Current It is usually given at a special temperature, usually 25°C, (77°F) and refers to the maximum amount of average current that can be permitted to flow in the forward direction. If this rating exceeds its limit, then the structure breakdown can occur.
- (e) **Repetitive Peak Forward Current** It is the maximum peak current that can be permitted to flow in the forward direction in the form of recurring pulses. The limiting value of this current is 450 mA.
- (f) Maximum Surge Current It is the maximum current permitted to flow in the forward direction in the form of nonrecurring pulses. The current should not equal this value for more than a few milliseconds.

The above diode rating are subject to change with temperature variations. If the operating temperature is more than that stated for the rating, then the ratings must be decreased.

2.19 ENERGY-BAND STRUCTURE OF OPEN-CIRCUITED PN JUNCTION

Consider that a *PN* junction has *P*-type and *N*-type materials in close physical contact at the junction on an atomic scale. Hence, the energy band diagrams of these two regions undergo relative shift to equalise the Fermi level. The Fermi level E_F should be constant throughout the specimen at equilibrium. The distribution of electrons or holes in allowed energy states is dependent on the position of the Fermi level. If this is not so, electrons on one side of the junction would have an average energy higher than those on the other side, and this causes transfer of electrons and energy until the Fermi levels on the two sides get equalised. However, such a shift does not disturb the relative position of the conduction band, valence band and Fermi level in any region. Equalisation of Fermi levels in the *P* and *N* materials of a *PN* junction is similar to equalisation of levels of water in two containers on being joined together.

The energy band diagram for a *PN* junction is shown in Fig. 2.24, where the Fermi level E_F is closer to the conduction band edge E_{cn} in the *N*-type material while it is closer to the valence band edge E_{vp} in the *P*-type material. It is clear that the conduction band edge E_{cp} in the *P*-type material is higher than the conduction band edge E_{cn} in the *N*-type material. Similarly, the valence band edge E_{vp} in the *P*-type material is higher than the conduction band edge E_{vn} in the *N*-type material. Similarly, the valence band edge E_{vp} in the *P*-type material is higher than the valence band edge E_{vn} in the *N*-type material. As illustrated in Fig. 2.24, E_1 and E_2 indicate the shifts in the Fermi level from the intrinsic conditions in the *P* and *N* materials respectively. Then the total shift in the energy level E_0 is given by

$$E_0 = E_1 + E_2 = E_{cp} - E_{cn} = E_{vp} - E_{vn}$$
(2.35)

This energy E_0 (in eV) is the potential energy of the electrons at the *PN* junction, and is equal to qV_0 , where V_0 is the contact potential (in volts) or contact difference of potential or the barrier potential.



Contact difference of potential A contact difference of potential exists across an open circuited *PN* junction. We now proceed to obtain an expression for E_0 . From Fig. 2.24, we find that

$$E_F - E_{vp} = \frac{1}{2} E_G - E_1 \tag{2.36}$$

$$E_{cn} - E_F = \frac{1}{2} E_G - E_2 \tag{2.37}$$

Combining Eqs (2.36) and (2.37), we get

$$E_0 = E_1 + E_2 = E_G - (E_{cn} - E_F) - (E_F - E_{vp})$$
(2.38)

We know that

 $np = N_C N_V e^{-E_G/kT}$ $np = n_i^2$ (Mass-action law)

and

From the above equations, we get

$$E_G = kT \ln \frac{N_C N_V}{n_i^2}$$
(239)

We know that for *N*-type material $E_F = E_C - kT \ln \frac{N_C}{N_D}$. Therefore, from this equation, we get

$$E_{cn} - E_F = kT \ln \frac{N_C}{n_n} = kT \ln \frac{N_C}{N_D}$$
(2.40)

Similarly for *P*-type material $E_F = E_V + kT \ln \frac{N_V}{N_A}$ Therefore, from this equation, we get

$$E_{F} - E_{vp} = kT \ln \frac{N_{V}}{P_{p}} = kT \ln \frac{N_{V}}{N_{A}}$$
(2.41)

Substituting from Eqs (2.39), (2.40) and (2.41) into Eq. (2.38), we get

$$E_{0} = kT \left[\ln \frac{N_{C} N_{V}}{n_{i}^{2}} - \ln \frac{N_{C}}{N_{D}} - \ln \frac{N_{V}}{N_{A}} \right]$$
$$= kT \ln \left[\frac{N_{C} N_{V}}{n_{i}^{2}} \times \frac{N_{D}}{N_{C}} \times \frac{N_{A}}{N_{V}} \right]$$
$$= kT \ln \frac{N_{D} N_{A}}{n_{i}^{2}}$$
(2.42)

As $E_0 = qV_o$, then the contact difference of potential or barrier voltage is given by

$$V_o = \frac{kT}{q} \ln \frac{N_D N_A}{n_i^2}$$
(2.43)

In the above equations, E_0 in electron volts and k is in electron volt per degree Kelvin. The contact difference of potential V_o is expressed in volt and is numerically equal to E_0 . From Eq. (2.42), we note that E_o (hence V_o) depends upon the equilibrium concentrations and not on the charge density in the transition region.

An alternative expression for E_0 may be obtained by substituting the equations of $n_n \approx N_D$, $p_n = \frac{n_i^2}{N_D}$, $n_n p_p = n_i^2$, $p_p \approx N_A$ and $n_p = \frac{n_i^2}{N_A}$ into Eq. (2.42). Then we get

$$E_0 = kT \ln \frac{p_{po}}{p_{no}} = kT \ln \frac{n_{no}}{n_{po}}$$
(2.44)

where subscript 0 represents the thermal equilibrium condition.

2.20 DIODE CURRENT EQUATION

The diode current equation relating the voltage V and current I is given by

$$I = I_o \left[e^{(V/\eta V_T)} - 1 \right]$$

where I = diode current

 I_o = diode reverse saturation current at room temperature

V = external voltage applied to the diode

 η = a constant, 1 for germanium and 2 for silicon

 $V_T = kT/q = T/11600$, volt-equivalent of temperature, i.e., thermal voltage

where $k = \text{Boltzmann's constant} (1.38 \times 10^{-3} \text{ J/K})$

q = charge of the electron (1.602 × 10⁻¹⁹ C)

T = temperature of the diode junction (K) = (°C + 273°)

At room temperature, (T = 300 K), $V_T = 26$ mV. Substituting this value in the current equation, we get

$$I = I_o[e^{(40 V/\eta)} - 1]$$

Therefore, for germanium diode, $I = I_o [e^{40V} - 1]$, since $\eta = 1$ for germanium. For silicon diode, $I = I_o [e^{20V} - 1]$, since $\eta = 2$ for silicon.

If the value of applied voltage is greater than unity, then the equation of diode current for germanium,

$$I = I_o (e^{40V})$$
$$I = I_o (e^{20V})$$

and for silicon,

When the diode is reverse biased, its current equation may be obtained by changing the sign of the applied voltage *V*. Thus, the diode current with reverse bias is

$$I = I_0 \left[e^{(-V/\eta V_T)} - 1 \right]$$

If $V >> V_T$, then the term $e^{(-V/\eta V_T)} \ll 1$, therefore $I \approx -I_o$, termed as reverse saturation current, which is valid as long as the external voltage is below the breakdown value.

2.21 ZENER AND AVALANCHE BREAKDOWN IN PN JUNCTIONS

The diode equation predicts that, under reverse bias conditions, a small constant current, the saturation current, I_o , flows due to minority carriers, which is independent of the magnitude of the bias voltage. But this prediction is not entirely true in practical diodes. There is a gradual increase of reverse current with increasing bias due to the ohmic leakage currents around the surface of the junction. Also, there is a sudden increase in reverse current due to some sort of *breakdown*, when the reverse bias voltage approaches a particular value called breakdown voltage, V_{BD} , as shown in Fig. 2.25. Once breakdown occurs, the diode is no longer blocking current and the diode current can be controlled only by the resistance of the external circuit.



Fig. 2.25 Breakdown in PN junction diodes

The breakdown occurs due to avalanche effect in which thermally generated minority carriers cross the depletion region and acquire sufficient kinetic energy from the applied potential to produce new carriers by removing valence electrons from their bonds. These new carriers will in turn collide with other atoms and will increase the number of electrons and holes available for conduction. This multiplication effect of free carriers may be represented by the following equation:

$$M = \frac{1}{1 - \left(\frac{V}{V_{BD}}\right)^n}$$

where M = carrier multiplication factor, which is the ratio of the total number of electrons leaving the depletion region to the number entering the region

V = applied reverse voltage

- V_{BD} = reverse breakdown voltage
- n = empirical constant, which depends on the lattice material and the carrier type, for *N*-type silicon, $n \approx 4$ and for *P*-type, $n \approx 2$

It is evident from the above equation that M is being very small for $V = 0.9 V_{BD}$. But when $V > 0.9 V_{BD}$, M is very large and the resulting reverse current is given by $I_R = MI_o$, where I_o is the reverse saturation current before breakdown.

As V approaches the breakdown voltage V_{BD} , the value of M will become infinite and there is a rapid increase in carrier density and a corresponding increase in current. Because of the cumulative increase in carrier density after each collision, the process is known as *avalanche breakdown*.

Even if the initially available carriers do not gain enough energy to disrupt bonds, it is possible to initiate breakdown through a direct rupture of the bonds because of the existence of strong electric field. Under these circumstances, the breakdown is referred to as *Zener breakdown*.

Voltage reference diodes that utilise the almost constant voltage characteristics in the breakdown region are also called *avalanche diodes* or sometimes *Zener diodes*. The Zener effect is in diodes with breakdown voltage below 6 V. The operating voltage in avalanche breakdown are from several volts to several hundred volts with power ratings up to 50 W.

True Zener diode action displays a negative temperature coefficient, i.e., breakdown voltage decreases with increasing temperature. True avalanche diode action exhibits a positive temperature coefficient, i.e., breakdown voltage increases with increasing temperature.

It is clear that the breakdown voltage for a particular diode can be controlled during manufacture by altering the doping levels in the junction. The breakdown voltage for silicon diodes can be made to occur at a voltage as low as 5 V with 10^{17} impurity atoms per cubic cm or as high as 1000 V when doped to a level of only 10^{14} impurity atoms per cubic cm.

2.22 HALL EFFECT AND DEVICES

When a transverse magnetic field B is applied to a specimen (thin strip of metal or semiconductor) carrying current I, an electric field E is induced in the direction perpendicular to both I and B. This phenomenon is known as the *Hall effect*.

A Hall-effect measurement experimentally confirms the validity of the concept that it is possible for two independent types of charge carriers, electrons and holes, to exist in a semiconductor.

The schematic arrangement of the semiconductor, the magnetic field and the current flow pertaining to the Hall effect are shown in Fig. 2.26. Under the equilibrium condition, the electric field intensity, E, due to the Hall effect must exert a force on the carrier of charge, q, which just balances the magnetic force, i.e.,

$$qE = Bqv_d$$

where v_d is the drift velocity. Also, the electric field intensity due to Hall effect is

$$E = \frac{V_H}{d}$$

where d is the distance between surfaces 1 and 2, and V_H is the Hall voltage appearing between surfaces 1 and 2. In an N-type semiconductor, the current is carried by electrons and these electrons will be forced downward towards side 1 which becomes negatively charged with respect to side 2.



Fig. 2.26 Schematic arrangement to observe the Hall effect

The current density (J) is related to charge density (ρ) by

$$J = \rho v_d$$

Further, the current density (J) is related to current (I) by

$$J = \frac{I}{\text{Area}} = \frac{I}{wd}$$

where w is the width of the specimen in the direction of magnetic field (B).

Combining the above relations, we get

$$V_H = Ed = B v_d d = \frac{BJd}{\rho} = \frac{BI}{\rho w}$$

The Hall coefficient, R_H , is defined by

$$R_H = \frac{1}{\rho}$$

so that $V_H = \frac{R_H}{w} BI$. A measurement of the Hall coefficient R_H determines not only the sign of the charge carriers but also their concentration. The Hall coefficient for a *P*-type semiconductor is positive, whereas it

is a negative for an *N*-type semiconductor. This is true because the Hall voltage in a *P*-type semiconductor is of opposite polarity to that in an *N*-type semiconductor.

The advantage of Hall-effect transducers is that they are non-contact devices with high resolution and small size.

Applications Hall effect is generally applied in semiconductor industry for material characterization. The Hall effect is used to find whether a semiconductor is *N*- or *P*-type and to determine the carrier concentration. If the terminal 2 becomes charged positively with respect to terminal 1, the semiconductor must be *N*-type and $\rho = nq$, where *n* is the electron concentration. On the other hand, if the polarity of V_H is positive at terminal 1 with respect to terminal 2, the semiconductor must be *P*-type and $\rho = pq$, where *p* is the hole concentration.

The mobility (μ) can also be calculated with simultaneous measurement of the conductivity (σ). The conductivity and the mobility are related by the equation $\sigma = \rho\mu$ or $\mu = \sigma R_H$.

Therefore, the conductivity for *N*-type semiconductor is $\sigma = nq\mu_n$ and for *P*-type semiconductor, $\sigma = pq\mu_p$, where μ_n is the electron mobility and μ_p is the hole mobility.

Thus, if the conductivity of a semiconductor is also measured along with R_H , then mobility can be determined from the following relations.

For *N*-type semiconductor, $\mu_n = \frac{\sigma}{nq} = \sigma R_H$ and for *P*-type semiconductor, $\mu_p = \frac{\sigma}{nq} = \sigma R_H$

Since V_H is proportional to *B* for a given current *I*, Hall effect can be used to measure the ac power and the strength of magnetic field and sense the angular position of static magnetic fields in a magnetic field meter. It is also used in an instrument called Hall-effect multiplier which gives the output proportional to the product

of two input signals. If *I* is made proportional to one of the inputs and *B* is made proportional to the second signal, then from the equation, $V_H = \frac{BI}{\rho_W}$, V_H will be proportional to the product of two inputs. Hall devices for such applications are made from a thin wafer or film of indium antimonide (InSb) or indium arsenide. As the material has a very high electron mobility, it has high Hall coefficient and high sensitivity.

An electrical current can be controlled by a magnetic field because the magnetic field changes the resistances of some elements with which it comes in contact. In the magnetic bubble memory, while read-out, the Hall effect element is passed over the bubble. Hence, a change in current of the circuit will create, say, a *one*. If there is no bubble, there will be a *zero* and there will be no current change in the output circuit. The read-in device would have an opposite effect, wherein the Hall device creates a magnetic field when supplied with a pulse of current. This, in turn, creates a little domain and then a magnetic bubble is created.

Some of the other applications of Hall-effect in industry are in measurement of velocity, rpm, sorting, limit sensing, non-contact current measurements and power measurements.

2.22.1 Hall-Effect Sensors

Hall effect sensors are used as devices for sensing magnetic field, for sorting magnetic material, as rpm meters. etc.

For application as a magnetic sensor, a current is driven through a conductor or doped semiconductor by applying appropriate bias voltage, and the Hall voltage is monitored. In the absence of a magnetic field, the

hall voltage will be zero. When the chip is subjected to a magnetic field in the correct direction, a hall voltage will appear indicating the presence of the magnetic field. This sensor can also be used to sort magnetic materials from others. If the magnetic field to be sensed is generated by an electromagnet, then same device serves as a current sensor or electronic multiplier with the hall voltage being proportional to the product of the two currents.

If, on the other hand, the semiconductor chip is placed between two magnetic pole pieces and simultaneously, a current is also passed in a direction perpendicular to the magnetic field, then a hall voltage will appear between the output pins. A small gap is created between one of the pole pieces and the chip which is accessible externally. If a ferrous wheel or a vane with ferromagnetic tip is inserted into the gap, the magnetic flux lines are cut and hall voltage drops to zero. If vane is circular and rotated at fixed speed, this gives you a timing signal. In a tachometer, the signal is used to measure rpm, whereas in automobile industry, timing is used for timing fuel injection and ignition.

2.22.2 Hall-Effect Wattmeter

Hall effect can be used to implement a wattmeter to measure electrical power dissipated in a load. The schematic sketch of the Hall-effect wattmeter is shown in Fig. 2.27(a), where the Hall-effect sample is typically a semiconductor material (usually Si).



Fig. 2.27 Wattmeter based on the Hall effect (a) Hall-effect wattmeter and (b) Flow of I_L through current coils

The load current I_L passes through two coils, which are called current coils and are shown as C in Fig. 2.27(b). These coils set up a magnetic field B_z such that $B_z \alpha I_L$. The Hall-effect sample is positioned in this field between the coils. The voltage V_L across the load drives a current $I_x = V_L/R$ through the sample, where R is a series resistance that is much larger than the resistance of the sample and that of the load. Normally, the current I_x , is very small and negligible compared to the load current. If d is the width of the sample, then the measured Hall voltage is

$$V_H = dE_H = dR_H J_x B_z \alpha I_x B_z \alpha V_L I_L$$

which is the electrical power dissipated in the load. The voltmeter that measures V_H can be calibrated to read directly the power dissipated in the load.

EXAMPLE 2.40

A semiconducting crystal 12 mm long, 5 mm wide and 1 mm thick has a magnetic flux density of 0.5 Wb/m^2 applied from front to back perpendicular to the largest faces. When a current of 20 mA flows lengthwise through the specimen, the voltage measured across its width is found to be 37 μ V. What is the Hall coefficient of the semiconductor? (AU Dec 2001, June 2012)

Solution Given t = 1 mm, $B = 0.5 \text{ Wb/m}^2$, I = 20 mA and $V_H = 37 \mu \text{V}$.

We know that Hall voltage, $V_H = R_H \frac{IB}{t}$

Therefore, the Hall coefficient is $R_H = \frac{V_H t}{IB} = \frac{37 \times 10^{-6} \times 1 \times 10^{-3}}{20 \times 10^{-3} \times 0.5} = 3.7 \times 10^{-6} \text{ m}^3 / \text{C}$

EXAMPLE 2.41

The Hall coefficient of certain silicon specimen was found to be $R_H = -7.35 \times 10^{-5} \text{ m}^3/\text{C}$ from 100 to 400 K. Determine the nature of the semiconductor. If the conductivity was found to be 200 ($\Omega \cdot \text{m}$)⁻¹, then calculate the density and mobility of the charge carrier. (AU Nov 2002, June 2012)

Solution Given $R_H = -7.35 \times 10^{-5} \text{ m}^3/\text{C}$ and $\sigma = 20 (\Omega \cdot \text{m})^{-1}$. Since the Hall coefficient is negative value, the given silicon specimen is *N*-type.

Hall coefficient, $R_H = -\frac{1}{nq}$, where the charge of an electron, $q = 1.602 \times 10^{-19}$ C

Therefore, the electron density is

$$n = -\frac{1}{R_H q} = \frac{1}{7.35 \times 10^{-5} \times 1.602 \times 10^{-19}} = 8.455 \times 10^{22} \,\mathrm{m}^{-3}$$

We know that $\sigma = nq\mu_e + pq\mu_h$

As the given semiconductor is *N*-type, p = 0

Hence, the electron mobility is

$$\mu_e = \frac{\sigma}{nq} = \frac{200}{8.455 \times 10^{22} \times 1.602 \times 10^{-19}} = 0.015 \,\mathrm{m}^2/\mathrm{V} \cdot \mathrm{s}$$

EXAMPLE 2.42

The Hall coefficient of a specimen of dopant silicon found to be $3.6 \times 10^{-4} \text{ m}^3 / \text{ C}$. The resistivity of the specimen is $8.93 \times 10^{-3} \Omega \cdot \text{m}$. Find the mobility and density of the charge carriers.

(AU April 2004, Dec 2010)

Solution Given $R_H = 3.6 \times 10^{-4} \text{ m}^3 / \text{C}$ and $\rho = 8.93 \times 10^{-3} \text{ W} \cdot \Omega \cdot \text{m}$. Since the Hall coefficient is positive value, the dopant silicon is *P*-type.

Hall coefficient, $R_H = \frac{1}{pa}$

Therefore, the density of the charge carriers i.e., holes is

$$p = \frac{1}{R_H q} = \frac{1}{3.66 \times 10^{-4} \times 1.602 \times 10^{-19}} = 1.7 \times 10^{22} \,\mathrm{m}^{-3}$$

We know that

$$\sigma = \frac{1}{\rho} = pq\mu_p$$

Hence, the mobility density of the holes is

$$\mu_p = \frac{1}{\rho pq} = \frac{1}{8.93 \times 10^{-3} \times 1.7 \times 10^{22} \times 1.602 \times 10^{-19}} = 0.041 \,\mathrm{m^2/V \cdot s}$$

EXAMPLE 2.43

An *N*-type semiconductor has Hall coefficient magnitude, $R_H = 4.16 \times 10^{-4} \text{ m}^3/\text{C}$. The conductivity of the semiconductor is 108 $(\Omega \cdot \text{m})^{-1}$. Calculate its charge carrier density and electron mobility at room temperature. (AU April 2003)

Solution Given $|R_H| = 4.16 \times 10^{-4} \text{ m}^3 / \text{C}$ and $\sigma = 108 \ (\Omega \cdot \text{m})^{-1}$.

Hall coefficient, $R_H = \frac{1}{nq}$, where the charge of an electron, $q = 1.602 \times 10^{-19} \text{ C}$

Therefore, the electron density is

$$n = \frac{1}{R_H q} = \frac{1}{4.16 \times 10^{-4} \times 1.602 \times 10^{-19}} = 1.5 \times 10^{22} \text{ m}^{-3}$$

We know that

$$\sigma = nq\mu_e + pq\mu_h$$

As the given semiconductor is *N*-type, p = 0

Hence, the electron mobility is

$$\mu_e = \frac{\sigma}{nq} = \frac{108}{1.5 \times 10^{22} \times 1.602 \times 10^{-19}} = 0.045 \text{ m}^2/\text{V} \cdot \text{s}$$

EXAMPLE 2.44

A copper strip 2 cm wide and 1 mm thick is placed in a magnetic field with B = 1.5 Wb/m². If a current of 200 A is set up in the strip, calculate Hall voltage that appears across the strip. Assume $R_H = 6 \times 10^{-7}$ m³/C.

Solution Given t = 1 mm, $B = 1.5 \text{ Wb/m}^2$, I = 200 A and $R_H = 6 \times 10^{-7} \text{ m}^3/\text{C}$.

Hall voltage,

$$V_H = R_H \frac{IB}{t} = 6 \times 10^{-7} \times \frac{200 \times 1.5}{10^{-3}} = 0.18 \text{ V}$$

EXAMPLE 2.45

An electric field of 100 V/m is applied to sample of *n*-type semiconductor whose Hall coefficient is -0.0125 m^3 / C. Determine the current density in the sample, assuming $\mu_e = 0.36 \text{ m}^2$ / V·s.

Solution Given $\mu_e = 0.36 \text{ m}^2 / \text{V} \cdot \text{s}$, E = 100 V/m and $R_H = -0.0125 \text{ m}^3 / \text{C}$.

Current density,

$$J = \frac{\mu_e E}{R_H} = \frac{0.36 \times 100}{-0.0125} = -2880 \text{ A/m}^2$$

EXAMPLE 2.46

In a Hall coefficient experiment, a current of 0.25 A is sent through a metal strip having thickness 0.2 mm and width 5 mm. The Hall voltage is found to be 0.15 mV when a magnetic field of 2000 gauss is used. Determine the carrier concentration and the drift velocity of the carriers.

Solution Given t = 0.2 mm, w = 0.5 mm, $V_H = 0.15$ mV, I = 0.25 A and B = 2000 gauss = 0.2 T (since 1 T = 10^4 gauss).

Carrier concentration,

$$n = \frac{IB}{V_H qt} = \frac{0.25 \times 0.2}{0.15 \times 10^{-3} \times 1.602 \times 10^{-19} \times 0.2 \times 10^{-3}} = 1.04 \times 10^{25} / \text{m}^3$$

Drift velocity of the carriers, $v_d = \frac{V_H}{wB} = \frac{0.15 \times 10^{-3}}{5 \times 10^{-3} \times 0.2} = 0.15 \text{ m/s}$

 V_H

EXAMPLE 2.47

A silicon plate is placed in the *X*-*Y* plane such that its length is parallel to *Y*-axis. A magnetic induction of 0.5 Wb/m² acts in the *z*-direction. The Hall coefficient is $R_H = 3.66 \times 10^{-4} \text{ m}^3 / \text{ C}$. The thickness, breath and length of the plate are 1 mm, 10 mm and 10 mm respectively. A current of 10^{-2} A flows through the plate in the *x*-direction. Calculate the voltage appearing across *y*-direction.

Solution Given
$$t = 1 \text{ mm}$$
, $B_2 = 0.5 \text{ Wb/m}^2$, $I_x = 10^{-2} \text{ A}$ and $R_H = 3.6 \times 10^{-4} \text{ m}^3 / \text{ C}$

Hall voltage,

$$=\frac{R_H I_x B_z}{t} = \frac{3.66 \times 10^{-4} \times 10^{-2} \times 0.5}{1 \times 10^{-3}} = 1.83 \times 10^{-3} \text{ V} = 1.83 \text{ mV}$$

EXAMPLE 2.48

In a Hall experiment, a current of 25 A is passed through a long foil of Ag 0.1 mm thick and 3 cm wide. If the magnetic field of flux density 0.14 Wb/m² is applied perpendicular to the foil, calculate the Hall voltage and mobility of e-s in Ag. The conductivity of Ag is $6.8 \times 10^7 (\Omega \cdot m)^{-1}$ and $R_H = -8.4 \times 10^{-11} \text{ m}^3/\text{C}.$

Solution Given t = 0.1 mm, $B = 0.14 \text{ Wb/m}^2$, I = 25 A and $R_H = -8.4 \times 10^{-11} \text{ m}^3/\text{C}$.

Hall voltage,

$$V_H = \frac{R_H IB}{t} = \frac{-8.4 \times 10^{-11} \times 25 \times 0.14}{0.1 \times 10^{-3}} = -2.94 \times 10^{-6} \,\mathrm{V} = -2.94 \,\mathrm{\mu V}$$

Mobility of electrons, $\mu_e = |\sigma R_H| = 6.8 \times 10^7 \times 8.4 \times 10^{-11} = 57.12 \times 10^{-4} \text{ m}^2/\text{V} \cdot \text{s}$

EXAMPLE 2.49

A current of 50 A is established in a slab of Cu 0.5 cm thick and 2 cm wide. The slab is placed in a field of 1.5 T. The magnetic field is perpendicular to the plane of the slab and to the current. The free electron concentration in Cu is 8.48×10^{28} electrons/m³. What is the Hall voltage across the width of the slab?

Solution Given t = 0.5 cm, B = 1.5 T = 1.5 Wb/m², I = 50 A and $n = 8.48 \times 10^{28}$ electrons/m³.

Hall voltage,

$$V_{H} = \frac{R_{H}IB}{t} = \frac{IB}{(nq) \times t} \qquad (\text{since } R_{H} = 1/nq)$$
$$= \frac{50 \times 1.5}{8.4 \times 10^{28} \times 1.602 \times 10^{-19} \times 0.5 \times 10^{-2}} = 1.116 \,\mu\text{V}$$

EXAMPLE 2.50

An *N*-type semiconductor has a Hall coefficient of 200 cm^3/C and its conductivity is 10 S/m. Find its electron mobility.

Solution Given $R_H = 200 \text{ cm}^3/\text{C}$ and $\sigma = 10 \text{ S/m}$.

Therefore, the electron mobility, $\mu_n = \sigma R_H = 10 \times 200 = 2000 \text{ cm}^2/\text{V-s}$

EXAMPLE 2.51

The conductivity of an *N*-type semiconductor is 10 S/m and its electron mobility is $50 \times 10^{-4} \text{ m}^2/\text{V-s}$. Determine the electron concentration.

Solution Given $\sigma = 10$ s/m and $\mu_n = 50 \times 10^{-4}$ m²/V-s.

We know that the electron mobility, $\mu_n = \frac{\sigma}{nq}$

Therefore, the electron concentration,

$$n = \frac{\sigma}{\mu q} = \frac{10}{50 \times 10^{-4} \times 1.6 \times 10^{-19}} = 12.5 \times 10^{21} \,\mathrm{m}^{-3}$$

EXAMPLE 2.52

A current of 20 A is passed through a thin metal strip, which is subjected to a magnetic flux density of 1.2 Wb/m². The magnetic field is directed perpendicular to the current. The thickness of the strip in the direction of the magnetic field is 0.5 mm. The Hall voltage is 60 V. Find the electron density.

Solution Given: $I = 20 \text{ A}, B = 1.2 \text{ Wb/m}^2, V_H = 60 \text{ V} \text{ and } w = 0.5 \text{ mm}$

We know that the number of conduction electrons, i.e., electron density,

$$n = \frac{BI}{V_H q w} = \frac{1.2 \times 20}{60 \times 1.6 \times 10^{-19} \times 0.5 \times 10^{-3}} = 5 \times 10^{21} \text{ m}^3$$

2.23 METAL-SEMICONDUCTOR JUNCTIONS

Metal-semiconductor junctions are very common in all semiconductor devices and have very high importance. Depending upon the doping concentration, materials, and the characteristics of the interface, the metal-semiconductor junctions can act as either an ohmic contact or as a Schottky barrier. An analysis of metal-semiconductor junction is presented in this section.

2.23.1 Structure of Metal-Semiconductor Junction

A metal-semiconductor junction, as the name indicates, consists of a metal in contact with a piece of semiconductor. The structure of a typical metal-semiconductor junction is shown in Fig. 2.28. The active junction is the interface between the metal, which acts as an anode, and the semiconductor. The other interface between the semiconductor and the metal, which acts as a cathode, is an Ohmic contact and there is no potential drop at this junction.



Fig. 2.28 Structure of a metal-semiconductor junction

2.23.2 Energy Band Diagram

The energy band diagram helps in identifying the barrier between the metal and the semiconductor. In order to understand the energy band structure at a metal–semiconductor junction, first let us consider the energy bands in metal and semiconductors separately, as shown in Fig. 2.29(a). The energy bands are aligned at the same vacuum level. When the metal and semiconductor are brought together, the Fermi levels do align themselves at thermal equilibrium. The condition that exists just before the thermal equilibrium is reached is depicted in Fig. 2.29(b).



Fig. 2.29 Energy band diagram of metal and semiconductor (a) before and (b) after contact is made

Let us define, Φ_{B} , the barrier height as the potential difference between the Fermi level of the metal and the band edge where the majority carriers exist. For an *N*-type semiconductor, the barrier height is given by the difference between the metal work function (Φ_{M}) and the electron affinity (χ).

Semiconductor Physics 2.69

$$\Phi_{BN} = \Phi_M - \chi \tag{2.45}$$

The work function, Φ_M varies depending upon surface preparation. For *P*-type semiconductor, the barrier height is given by the difference between the valence band level and the Fermi level in the metal,

$$\Phi_{BP} = \chi + \frac{E_g}{q} - \Phi_M \tag{2.46}$$

where E_g is the energy gap between the conduction and valence bands. The sum of the barrier heights on *N*-type and *P*-type substrate is expected to be equal to the energy gap, E_g , i.e., $(\Phi_{BN} + \Phi_{BP}) q = E_g$.

In a metal-semiconductor junction, a barrier is formed if the Fermi level of the metal is somewhere between the valence and conduction band edges of the semiconductor, as shown in Fig. 2.29(b). Let us also define a built-in potential (Φ_1) as the difference between the Fermi level of the metal and the Fermi level of the semiconductor.

For an N-type semiconductor, the barrier height is given by

$$\Phi_{BN} = \Phi_M - \chi$$

$$\Phi_{IN} = \Phi_{BN} - \frac{E_C - E_F}{q} = \Phi_M - \chi - \frac{E_C - E_F}{q}$$
(2.47)

For a *P*-type semiconductor, the Fermi level is closer to the valence band and the built-in potential is given by

$$\Phi_{IP} = \chi + \frac{E_F - E_V}{q} - \Phi_M \tag{2.48}$$

The Fermi level in an N-type semiconductor is given by

$$E_F = E_C - kT \ln \frac{N_C}{N_D} \tag{2.49}$$

and the Fermi level in a P-type semiconductor is given by

$$E_F = E_V + kT \ln \frac{N_V}{N_A} \tag{2.50}$$

Substitution Eqns. (2.49) and (2.50) in Eqns. (2.47) and (2.48), respectively, would give expressions for built in potentials in terms of the barrier height and doping concentration, as follows.

$$\Phi_{IN} = \Phi_{BN} - \frac{E_C - E_F}{q} = \Phi_{BN} - \frac{kT}{q} \ln \frac{N_C}{N_D} \text{ for } N \text{-type semiconductor}$$
(2.51)

and

$$\Phi_{IP} = \Phi_{BP} - \frac{E_F - E_V}{q} = \Phi_{BP} - \frac{kT}{q} \ln \frac{N_V}{N_A} \text{for } P \text{-type semiconductor}$$
(2.52)

2.23.3 Thermal Equilibrium

After the metal and semiconductor have been brought into contact, electrons start to flow from the semiconductor into the metal, and as a result, a depletion region of width x_d , with uncompensated donors (positive charge) is formed. Electrons continue to flow into the metal until the Fermi energy levels of metal and semiconductor align with each other. In metal, the electron current forms a negative surface charge layer. This results in an electric field and the band edges are lowered in the semiconductor as shown in Fig. 2.30.



Fig. 2.30 Energy band diagram in thermal equilibrium

Table 2.3 gives the work function (F_M) and electron affinity (χ) of some commonly used metals and semiconductors.

Table 2.3 Work function and electron affinity of some metals and semiconductors

Element	Work function, $\Phi_M(V)$
Aluminium, Al	4.28
Silver, Ag	4.26
Gold, Au	5.1
Chromium, Cr	4.5
Nickel, Ni	5.15
Platinum, Pt	5.65
Titanium, Ti	4.33
Tungsten, W	4.55
	<i>Electron affinity</i> , χ (Volt)
Silicon, Si	4.01
Germanium, Ge	4.13
Gallium Arsenide, GaAs	4.07

2.23.4 Forward and Reverse Bias

When an external bias is applied, the metal-to-semiconductor barrier remains unchanged, whereas, the semiconductor-to-metal barrier is either decreased (forward bias) or increased (reverse bias).

When the metal is connected to a positive bias with respect to the semiconductor. The Fermi energy level of the metal is lowered from its equilibrium level. The depletion region is narrowed, and the potential barrier in the semiconductor is reduced. The number of electrons that diffuse from semiconductor to metal is now

Semiconductor Physics 2.71

more than the number of electrons that drift from metal into the semiconductor. Thus, there will be a positive current through the device. Figure 2.31(a) illustrates a metal–semiconductor junction under forward-bias condition.

If the metal is connected to a negative bias with respect to the semiconductor, the metal is charged even more negatively than without any bias. The Fermi energy level of the metal is raised. The electrons in the semiconductor are repelled even more. The depletion region becomes wider and the potential barrier on semiconductor side is further increased, as shown in Fig. 2.31(b). However, the barrier on the metal side remains unchanged and limits the flow of electrons. A small current flows as a result of a few electrons in the metal acquiring enough thermal energy to overcome barrier.





Fig. 2.31(a) Energy band diagram under forward bias

Fig. 2.31(b) Energy band diagram under reverse bias

EXAMPLE 2.53

A metal–semiconductor junction is made of silver and silicon with $N_D = 4 \times 10^{17}$ cm⁻³. Calculate the barrier height and the built-in potential.

Solution The work function (Φ_M) and electron affinity (χ) for silver and silicon are 4.26 V and 4.01 V, respectively (from Table 2.3). The barrier height for *N*-type material, from Eq. (2.45), is given by $\Phi_{BN} = \Phi_M - \chi = 4.26 - 4.01 = 0.25 \text{ V}$

The built-in potential is given by Eq. (2.51), i.e.,

$$\Phi_{IN} = \Phi_{BN} - \frac{kT}{q} \ln \frac{N_C}{N_D} = 0.25 - \frac{1.38 \times 10^{-23} \times 300}{1.602 \times 10^{-19}} \ln \frac{2.8 \times 10^{25}}{4 \times 10^{17}} = 0.47 \text{ V}$$

(The density of state at conduction band edge for silicon, $N_C = 2.8 \times 10^{25}$ from Table 2.2)

2.24 OHMIC CONTACTS

An Ohmic contact is another type of metal–semiconductor junction. It is formed by applying a metal to a heavily doped semiconductor. Here, the current is conducted equally in both directions and there will be a very little voltage drop across the junction. The usage of Ohmic contacts is to connect one semiconductor device to another on an IC, or to connect an IC to its external terminals.

Ohmic contacts are very common in semiconductor devices. Metal-semiconductor contacts cannot be considered to offer a resistance as low as that of two metals connected to each other. Metal-semiconductor junctions can act as either a rectifying junction or an Ohmic contact depending upon the Fermi energy levels of the metal and the semiconductor used. A proper choice of metal and semiconductor can offer a low resistance Ohmic contact. Alternatively, contacts that have a thin barrier can be created by heavily doping the semiconductor through which the carriers can tunnel. Both these types of contacts are presented in this section.

A metal-semiconductor junction can be an Ohmic contact if the Schottky barrier height, Φ_B , is zero or negative. This means, for an *N*-type semiconductor, that the metal work function, Φ_M , is either close to or smaller than the electron affinity (χ) of the semiconductor; and for a *P*-type semiconductor, the metal work function is either close to or greater than the sum of electron affinity and the bandgap energy.

That is, $\Phi_M \leq \chi$ for an *N*-type semiconductor

or, $\Phi_M \ge \chi + E_g$ for a *P*-type semiconductor

Figure 2.32 illustrates the energy band diagram when Q_M is less than the electron affinity, χ . In this case, the electrons flow from the metal into the semiconductor thus forming a positive surface charge layer in metal. The resultant electric field sets up an electric potential and the energy bands of the semiconductor bend downward. There is no barrier for the flow of electrons in both directions. The current is directly proportional to the potential across the junction and is symmetric about the origin, as shown in Fig. 2.33.

A more practical method of providing contacts in semiconductor devices is to create a junction that consists

of a thin barrier. Such contacts are also referred to as **Fig. 2.32** Energy band diagram in an Ohmic contact tunnel contacts. Such contacts have a positive barrier at the junction and a heavy doping in the semiconductor. This creates a very thin barrier separating the metal from the semiconductor, through which the carriers can easily tunnel. Figure 2.34 shows the energy band in a heavily doped Ohmic contact.



Fig. 2.33 V-I characteristics of an Ohmic contact



 $q(\Phi_{a}-\gamma)$

ÅΕ



Fig. 2.34 Energy band levels in an Ohmic contact
2.25 SCHOTTKY DIODE

The metal contacts are required to be ohmic and no PN junctions to be formed between the metal and silicon layers. The N^+ diffusion region serves the purpose of generating ohmic contacts. On the other hand, if aluminium is deposited directly on the *N*-type silicon, then a metal-semiconductor diode can be said to be formed. Such a metal semiconductor diode junction exhibits the same type of *V*-*I* characteristics as that of an ordinary *PN* junction.

The cross-sectional view and symbol of a Schottky barrier diode are shown in Fig. 2.35(a) and (b) respectively. Contact 1 shown in Fig. 2.35(a) is a Schottky barrier and the contact 2 is an ohmic contact. The contact potential between the semiconductor and the metal generates a barrier for the flow of conducting electrons from semiconductor to metal. When the junction is forward biased, this barrier is lowered and the electron flow is allowed from semiconductor to metal, where the electrons are in large quantities.

The majority carriers carry the conduction current in the Schottky diode whereas in the *PN* junction diode, minority carriers carry the conduction current and it incurs an appreciable time delay from ON state to OFF state. This is due to the fact that the minority carriers stored in the junction have to be totally removed.



Fig. 2.35 Schottky barrier diode: (a) Cross-sectional view, (b) Symbol

This characteristic puts the Schottky barrier diode at an advantage since it exhibits negligible storage time to flow the electron from *N*-type silicon into aluminium almost right at the contact surface, where they mix with the free electrons. The other advantage of this type of diode is that it has less forward voltage (approximately 0.4 V). Thus, it can be used for clamping and detection in high frequency applications and microwave integrated circuits.

Schottky barrier diode is an extension of the oldest semiconductor device that is the point contact diode. Here, the metal–semiconductor interface is a surface, the Schottky barrier rather than a point contact. The Schottky diode is formed when a metal, such as Aluminum, is brought into contact with a moderately doped *N*-type semiconductor as shown in Fig. 2.36. It is a unipolar device because it has electrons as majority carriers on both sides of the junction. Hence, there is no depletion layer formed near the junction. It shares the advantage of point contact diode in that there is no significant current from the metal to the semiconductor

with reverse bias. Thus, the delay present in the junction diodes due to hole–electron recombination time is absent here. Hence, because of the large contact area between the metal and semiconductor than in the point contact diode, the forward resistance is lower, and so is noise.



Fig. 2.36 Schottky diode: (a) metal–semiconductor contact (b) circuit symbol

The forward current is dominated by electron flow from semiconductor to metal, and the reverse current is mainly due to electron flow from metal to semiconductor. As there is very little minority carrier injection from semiconductor into metal, Schottky diodes are also said

to be majority carrier devices.

The diode is also referred to as hot carrier diode because when it is forward biased, conduction of electrons on the Nside gains sufficient energy to cross the junction and enter the metal. Since these electrons plunge into the metal with large energy, they are commonly called as hot carriers.

Figure 2.37 shows the *V-I* characteristics of a Schottky diode and a *PN* junction diode. The current in a *PN* junction diode is controlled by the diffusion of minority carriers whereas the current in the Schottky diode results from the flow of majority carriers over the potential barrier at the metal-semiconductor junction. The reverse saturation current for a Schottky diode is larger than that of a *PN* junction diode. The storage time for



Fig. 2.37 V–I characteristics of Schottky barrier diode and PN junction diode

a Schottky diode is theoretically zero. The Schottky diode has a smaller turn-on voltage and shorter switching time than the *PN* junction diode.

Schottky diode can be used for rectification of signals of frequencies even exceeding 300 MHz. It is commonly used in switching power supplies at frequencies of 20 GHz. Its low noise figure finds application in sensitive communication receivers like radars. It is also used in clipping and clamping circuits and in computer gating.

2.26 TUNNEL DIODE

The Tunnel or Esaki diode is a thin-junction diode which exhibits negative resistance under low forward bias conditions.

An ordinary *PN* junction diode has an impurity concentration of about 1 part in 10^8 . With this amount of doping, the width of the depletion layer is of the order of 5 microns. This potential barrier restrains the flow of carriers from the majority carrier side to the minority carrier side. If the concentration of impurity atoms is greatly increased to the level of 1 part in 10^3 , the device characteristics are completely changed. The width of the junction barrier varies inversely as the square root of the impurity concentration and therefore, is reduced from 5 microns to less than 100 Å (10^{-8} m). This thickness is only about $1/50^{\text{th}}$ of the wavelength of visible light. For such thin potential energy barriers, the electrons will penetrate through the junction rather than surmounting them. This quantum mechanical behavior is referred to as tunneling and hence, these high-impurity-density *PN* junction devices are called tunnel diodes.

The V-I characteristic for a typical germanium tunnel diode is shown in Fig. 2.38. It is seen that at first, forward current rises sharply as applied voltage is increased, where it would have risen slowly for an ordinary PN junction diode (which is shown as dashed line for comparison). Also, reverse current is much larger for comparable back bias than in other diodes due to the thinness of the junction. The interesting portion of the characteristic starts at the point A on the curve, i.e., the peak voltage. As the forward bias is increased beyond this point, the forward current drops and continues to drop until point B is reached. This is the valley voltage. At B, the current starts to increase once again and does so very rapidly as bias is increased further. Beyond this point, characteristic resembles that of an ordinary diode. Apart from the peak voltage and valley voltage,

the other two parameters normally used to specify the diode behaviour are the peak current and the peak-tovalley current ratio, which are 2 mA and 10 respectively, as shown.



Fig. 2.38 V–I characteristic of tunnel diode

The V-I characteristic of the tunnel diode illustrates that it exhibits dynamic resistance between A and B. Figure 2.39 shows energy level diagrams of the tunnel diode for three interesting bias levels. The shaded areas show the energy states occupied by electrons in the valence band, whereas the cross hatched regions represent energy states in the conduction band occupied by the electrons. The levels to which the energy states are occupied by electrons on either side of the junctions are shown by dotted lines. When the bias is zero, these lines are at the same height. Unless energy is imparted to the electrons from some external source, the energy possessed by the electrons on the N-side of the junction is insufficient to permit to climb over the junction barrier to reach the P-side. However, quantum mechanics show that there is a finite probability for the electrons to tunnel through the junction to reach the other side, provided there are allowed empty energy states in the P-side of the junction at the same energy level. Hence, the forward current is zero.

When a small forward bias is applied to the junction, the energy level of the *P*-side is lower as compared with the *N*-side. As shown in Fig. 2.39(b), electrons in the conduction band of the *N*-side see empty energy level on the *P*-side. Hence, tunneling from *N*-side to *P*-side takes place. Tunneling in other directions is not possible because the valence band electrons on the *P*-side are now opposite to the forbidden energy gap on the *N*-side. The energy band diagram shown in Fig. 2.39(b), is for the peak of the diode characteristic.

When the forward bias is raised beyond this point, tunneling will decrease as shown in Fig. 2.39(c). The energy of the *P*-side is now depressed further, with the result that fewer conduction band electrons on the

N-side are opposite to the unoccupied *P*-side energy levels. As the bias is raised, forward current drops. This corresponds to the negative resistance region of the diode characteristic. As forward bias is raised still further, tunneling stops altogether and it behaves as a normal *PN* junction diode.



Fig. 2.39 Energy level diagrams of tunnel diode

2.26.1 Equivalent Circuit

The equivalent circuit of the tunnel diode when biased in the negative resistance region is as shown in Fig. 2.40(a). In the circuit, R_s is the series resistance and L_s is the series inductance which may be ignored except at highest frequencies. The resulting diode equivalent circuit is thus reduced to parallel combination of the junction capacitance C_j and the negative resistance $-R_n$. Typical values of the circuit components are $R_s = 6 \Omega$, $L_s = 0.1$ nH, $C_j = 0.6$ pF and $R_n = 75 \Omega$. The symbol of tunnel diode is shown in Fig. 2.40(b).



Fig. 2.40 (a) Equivalent circuit of tunnel diode, (b) Symbol of tunnel diode

Applications

- 1. Tunnel diode is used as an ultra-high speed switch with switching speed of the order of ns or ps
- 2. As logic memory storage device
- 3. As microwave oscillator
- 4. In relaxation oscillator circuit
- 5. As an amplifier

Advantages

- 1. Low noise
- 2. Ease of operation
- 3. High speed
- 4. Low power

Disadvantages

- 1. Voltage range over which it can be operated is less than 1 V.
- 2. Being a two terminal device, there is no isolation between the input and output circuit.

2.27 MOS CAPACITOR

The main part of the MOSFET is the Metal-Oxide-Semiconductor structure, called MOS capacitor. An MOS transistor is an MOS capacitor with two *PN* junctions adjoining the capacitor. The passive capacitors can be fabricated using MOS technology. The capacitance of a device is represented by

$$C = \frac{dQ}{dV}$$

where dQ is the magnitude of the differential change in the charge on one plate as a function of the differential change in voltage dV across the capacitor. Then the capacitance is measured as a function of the applied DC gate voltage. Any small variation in voltage across the MOS capacitor results in the corresponding variation in the charges on metal gate. Further, the hole accumulation charge also varies accordingly. Similar to a parallel plate capacitor, small variation in charge density takes place at the edges of oxide layer.

Construction

The MOS capacitor consists of a Metal-Oxide-Semiconductor structure as shown in Fig. 2.41. Here, the semiconductor *P*-type substrate which is the bottom electrode formed a N^+ layer with a thin oxide layer (SiO₂) and a top metal contact, referred to as the gate. An ohmic contact, called the bulk contact is formed by a second metal layer to the back of the semiconductor. This is known as *N*-type MOS or NMOS capacitor since the inversion layer contains electrons.





Initially, zero charge is trapped in the oxide and also, no charge is trapped at the oxide-semiconductor interface. The MOS capacitor acts as an enhancement mode device for all values of positive oxide charge. Also, the threshold voltage becomes more negative, as the charge increases.

There are three different operating modes in the MOS capacitor, namely accumulation, depletion and inversion.

- (i) The first operating condition, accumulation is below the flat-band voltage, V_B . Here, flat-band voltage is defined as the applied gate voltage such that there is no band bending in the semiconductor and hence the space charge is zero.
- (ii) The second mode of operation, depletion is in between the flat-band voltage, V_{FB} and the threshold voltage, V_T which is the gate voltage that must be applied to achieve the threshold inversion point, and
- (iii) The third operating condition, inversion is above the threshold voltage, V_T .

The charge distributions associated with these three modes are also shown in Fig. 2.42.



Fig. 2.42 Distribution of charge in an N-type MOS structure with P-type substrate under (a) accumulation, (b) depletion and (c) inversion conditions

The ideal low frequency capacitance versus gate voltage, i.e., *C*-*V* characteristics of the MOS capacitor with a *P*-type substrate is shown in Fig. 2.43.



Fig. 2.43 An ideal capacitance vs gate voltage of MOS capacitor with P-type substrate

Accumulation

Accumulation occurs when negative voltage is applied to the gate, which is less than the flat-band voltage, V_{FB} . Here, the holes are attracted by the negative charge from the substrate to the oxide-semiconductor interface. In order to have potential variation within the oxide, a small amount of band bending is required to build up the accumulation charge

The capacitance per unit area of the MOS capacitor for this accumulation is only the oxide capacitance, C_{ox} , i.e.,

$$C_{acc} = C_{ox} = \frac{\mathcal{E}_{ox}}{t_{ox}}$$

where ε_{ox} is the oxide permittivity and t_{ox} is the oxide thickness.

Depletion

Depletion occurs when the applied gate voltage is above the flat-band voltage, V_{FB} but below the threshold voltage, V_T . The positive charge on the gate pushes the mobile holes into the substrate. This leads to increase in negative charges in the semiconductor. This results in the depletion of mobile carriers at the interface and hence, a negative charge due to the ionized acceptor ions is left in the space charge region. This charge is mainly due to the depletion of the semiconductor starting from the oxide-semiconductor interface. Hence, there will be an increase in depletion layer width with increase in gate voltage.

The voltage separating the accumulation and depletion regions is known as the *flat-band voltage*, V_{FB} . The flat-band condition exists when no charge is present in the semiconductor so that the silicon energy band is flat.

Flat-band diagram

Flat-band voltage and the threshold voltage are the two important parameters of the MOS structure. The flat band voltage is the gate voltage which is applied to achieve the flat-band condition. Here, the conduction band, E_C and the valence band, E_V in the semiconductor do not bend as shown in Fig. 2.44 which shows an aluminum-silicon dioxidesilicon MOS structure. Here, a voltage, V_{FB} , must be applied to obtain this flat-band as indicated in the diagram. Also, $E_{F,M}$ is the work function of the aluminum gate, C_{ox} is the electron affinity of the oxide and E_G is the bandgap energy of silicon,. The bandgap energy of the oxide is between 8 and 9 eV.

When the applied gate voltage is equal to the work function difference between the gate metal and the semiconductor, the flat-band voltage is obtained. The expression for the flat-





band voltage can be modified based on the fixed charge in the oxide-silicon interface.

Here, the oxide capacitance, C_{ox} and the capacitance of the depletion region, C_{SD} are in series. Hence, the total capacitance in the depletion region becomes

$$\frac{1}{C_{\rm dep}} = \frac{1}{C_{\rm ox}} + \frac{1}{C_{\rm SD}}$$

 $C_{\rm dep} = \frac{C_{\rm ox}C_{\rm SD}}{C_{\rm ox} + C_{\rm SD}}$

Therefore,

Since $C_{\text{ox}} = \frac{\varepsilon_{\text{ox}}}{t_{\text{ox}}}$ and $C_{\text{SD}} = \frac{\varepsilon_s}{x_d}$, the above equation can be written as

$$C_{\rm dep} = \frac{C_{\rm ox}}{1 + \frac{C_{\rm ox}}{C_{\rm SD}}} = \frac{\varepsilon_{ox}}{t_{ox} + \left(\frac{\varepsilon_{ox}}{\varepsilon_{s}}\right) x_{d}}$$

where ε_s is the permittivity of the semiconductor, x_d is the space charge width and t_{ox} is the oxide thickness.

Due to increase in the space charge width, there will be a decrease in total capacitance C_{dep} . Further, if the depletion width reaches the maximum value, there will be zero inversion charge density. This condition leads to a minimum capacitance, which is given by

$$C_{\min} = \frac{\varepsilon_{\text{ox}}}{t_{ox} + \left(\frac{\varepsilon_{\text{ox}}}{\varepsilon_{s}}\right) x_{dT}}$$

where x_{dT} is the maximum space charge width.

Inversion

The formation of the inversion layer charge in the semiconductor adjacent to the oxide is an important characteristic of the MOS structure. Here, the inversion occurs at voltages beyond the threshold voltage, V_T . When the voltage across the semiconductor increases beyond twice the semiconductor bulk voltage, another type of negative charge emerges at the oxide-semiconductor interface. This charge is due to minority carriers which form an inversion layer. Further, increase in gate voltage leads to decrease in the depletion layer width as the charge in the inversion layer increases exponentially with the surface potential. Here, a negatively charged inversion layer occurs during inversion at the oxide-semiconductor interface in addition to the depletion-layer. The positive gate voltage attracts the minority charge carriers to the interface and this will result in the formation of inversion layer.

The energy band diagram of an NMOS capacitor biased in inversion is shown in Fig. 2.45. Here, the oxide layer is assumed as a semiconductor with large bandgap, which blocks the flow of charge carriers between semiconductor and gate metal. There will be a constant band bending in the semiconductor due to the presence of a depletion layer. So, the Fermi energy level is very near to the edge of conduction band at the semiconductor-oxide interface when there is high electron density. Even due to the applied gate voltage, the semiconductor will be in thermal equilibrium condition. Further, there will be an electric field due to the gate voltage and such fields cannot result in a non-equilibrium condition, as seen in a *PN* diode with zero bias.



If the inversion charge can respond to the change in capacitor voltage, then the capacitance in inversion region is again just the oxide capacitance.

i.e.,

$$C_{\rm inv} = C_{\rm ox} = \frac{\mathcal{E}_{\rm ox}}{t_{\rm ox}}$$

Referring to Fig. 2.43, the three regions of operation shown by dashed lines correspond to three components, C_{ox} , C_{SD} and C_{min} . The ideal total capacitance of the MOS capacitor is shown by solid line.

The flat-band condition occurs between the accumulation and depletion regions. Hence the capacitance at flat-band is given by

$$C_{\rm FB} = \frac{\varepsilon_{\rm ox}}{t_{\rm ox} + \left(\frac{\varepsilon_{\rm ox}}{\varepsilon_s}\right) \sqrt{\left(\frac{kT}{q}\right) \left(\frac{\varepsilon_s}{qN_A}\right)}}$$

where N_A is the acceptor doping concentration and q is the charge of an electron. Here, the flat-band capacitance C_B , is a function of oxide thickness t_{ox} and semiconductor doping N_A .

2.28 POWER TRANSISTOR

A Bipolar Junction Transistor (BJT) is a three-terminal semiconductor device in which the operation depends on the interaction of both majority and minority carriers and hence the name *bipolar*. The BJT is analogous to a vacuum triode and is comparatively smaller in size. It is used in amplifier and oscillator circuits, and as a switch in digital circuits. It has wide applications in computers, satellites and other modern communication systems.

The bipolar power transistor is a three layer NPN or PNP device. The BJT consists of a silicon (or germanium) crystal in which a thin layer of N-type silicon is sandwiched between two layers of P-type silicon. This transistor is referred to as PNP. Alternatively, in a NPN transistor, a layer of P-type material is sandwiched between two layers of N-type material. Since the collector current I_C is a function of the base current I_B , a change in the base current provides a corresponding amplified change in the collector current for a given collector emitter voltage V_{CE} . The ratio of the collector current I_C to the base current I_B is called current gain β i.e., $\beta = I_C/I_B$, which is in the order of 15 to 150.

Consider an *NPN* power transistor biased in common emitter configuration as shown in Fig 2.46(a). A family of collector characteristic curves are obtained when I_C vs V_{CE} is plotted for various values of I_B for Common Emitter configuration as shown in Fig. 2.46(b).



Fig. 2.46 Common Emitter configuration of NPN power transistor (a) Circuit diagram (b) Collector characteristics

When the base is maintained at 0.7V and the emitter and collector are at 0V, I_B flows through base-emitter junction due to low impedance path and I_C is zero. When both the junctions are forward biased, the transistor is in saturation region of operation.

When V_{CC} is increased, V_{CE} increases gradually because of the increase in the collector current. Ideally, when V_{CE} exceeds 0.7V, the base-collector junction is reverse biased and the transistor goes into the active region of its operation. Once the base-collector junction is reversed biased, I_C remains constant for a given value of I_B as V_{CE} continues to increase. Actually, I_C increases very slightly as V_{CE} increases due to widening of the base-collector depletion region.

When V_{CE} reaches a sufficiently high voltage, the reverse biased base-collector junction goes into breakdown and the collector current increases rapidly. This collector region is called breakover region. $V_{CE}(V)$

When $I_B = 0$, the transistor is in cut-off region even though there is a small collector leakage current as indicated. The transistor can act as open or close switch based on the base current, I_B . If $I_B = 0$, then there will be a very small leakage current of I_C , and it acts as an open switch. If I_B reaches to a value which can drive the transistor into saturation state, then the transistor acts as a closed switch.

In order to maintain control, I_B should be just enough to keep the device in saturation. When the transistor is initially turned-ON, the base current must be high so as to give a fast turn-ON. Any change in I_C should be matched by a change in I_B . When the transistor is turned-OFF, I_B must be decreased at a rate in such a way that I_C can follow so as to avoid secondary breakdown. In the OFF-state, a small reverse current I_B is maintained to avoid very high collector current.

The power loss in a transistor is the product of V_{CE} and I_C . As a switch, the power loss of transistor is small, because (i) in the open position, leakage current is small and (ii) in the close position, the saturation voltage is small.

To utilize the transistor in an effective manner without over-heating during switching, safe operation area characteristics, as shown in Fig. 2.47, can be utilized. When switching between the two states occurs, it is necessary that the voltage and current values should be within the rectangular area at all times during the switching period.

The largest instantaneous power loss can be controlled for longer switching times. The transistor switching loss can be high, because during the switching both the voltage across the transistor and current through



Fig. 2.47 Safe Operating Area Characteristics

it can be high. A high switching frequency can infer that predominant loss is due to switching. The exact switching loss is a function of the load circuit parameters and change in base current.

The current gain of a transistor can be improved by using Darlington arrangement in which the base drive current is obtained from another transistor. In this case, the overall current gain of 250 is possible but switching time should be made longer.

Practical Power Transistors

Physical transistors have certain limitations in terms of maximum current, voltage and power. These limitations are not considered for normal transistors, because it was assumed that the transistors which are capable of handling the current, voltage and power dissipation within a transistor does not cause any damage to the circuit.

But in the design of power amplifiers, it is necessary to consider the limitations of a transistor. The limitations are maximum rated current (amp), maximum rated voltage (volts) and maximum rated power (watts). As power amplifiers use BJTs and MOSFETs, the limitation effect is considered on these two transistors. The maximum power limitation is related to maximum allowed temperature of a transistor (BJTs or MOSFETs) which in turn is a function of the rate at which heat is removed.

Construction of power transistor

The structure of a vertically oriented *NPN* power transistor with the doping levels and thickness of the layers is shown in Fig. 2.48. This type of configuration has a large cross sectional area to handle large currents and minimize the thermal resistance of the transistor. Here, the collector terminal is at the bottom.



Fig. 2.48 Cross section of a vertical NPN power BJT

The primary collector N^- region, called drift region, has a low-doped impurity concentration 10^{20} m⁻³ in such a way that the voltage can be applied across base-collector terminals without initiating breakdown. This region has a thickness of about 50–200 µm. The thickness of the drift region determines the breakdown voltage of the transistor.

Another N^+ region has a higher doping concentration (typically 10^{25} m^{-3}) which reduces the collector resistance and makes contact with the external terminal. This region has a thickness of around 250 µm. The doping in the emitter layer is large (typically 10^{25} m^{-3}), whereas the base doping is comparatively less (10^{22} m^{-3}) .

A sufficient base width is required to prevent punch-through breakdown. Since the small base thickness of about $5-20 \ \mu m$ reduces the breakdown voltage, the amplification capabilities and the breakdown voltages are to be compromised in power transistors. A large base-collector voltage implies a relatively large space-charge width induced in the collector and base regions.

V-I Characteristics

Power transistors are generally large area devices. The properties of power transistors vary from small-signal devices in terms of differences in geometry and doping concentrations. Table 2.4 compares the various parameters of small-signal BJT to those of two power BJTs. The current gain of power transistors is in the range of 5 to 20 which is smaller to that of small-signal BJTs. But the current gain is a strong function of collector current and temperature. The current gain versus collector current characteristics of 2N3055 power BJT at various temperatures is shown in Fig. 2.49. The current gain drops off for high current levels. The parasitic resistances in the base and collector regions may become significant by affecting the transistor terminal characteristics.

Parameter	Small-Signal BJT (2N2222A)	Power BJT (2N3055)	Power BJT (2N6078)
V _{CE(max)} in V	40	60	250
$I_{C(\max)}$ in A	0.8	15	7
$P_{D(\max)}$ in W (at T = 25°C)	1.2	115	45
β	35-100	5–20	12–70
f_T in MHz	300	0.8	1

 Table 2.4
 Parametric comparison of small-signal BJT with power BJTs



Fig. 2.49 *DC* characteristics of β vs I_c

The maximum rated collector current $I_{c,\text{rated}}$ is related to the maximum current that the wires connecting the semiconductor to the external terminals can handle or the collector current at which the current gain falls below a minimum specified value or the current that leads to the maximum power dissipation when the transistor is in saturation.

In BJTs, the maximum voltage limitation is associated with avalanche breakdown in the reverse-biased base-collector junction. In common-emitter configuration, the breakdown voltage mechanism involves the transistor gain and breakdown phenomenon on the *PN* junction. Typical I_C versus V_{CE} characteristics curve of BJT is shown in Fig. 2.50. When the base terminal is open circuited, i.e., $I_B = 0$, the breakdown voltage is V_{CE0} and from Fig. 2.50, its value is 130 V.



When the transistor is biased in the active region, the collector current begins to increase significantly before breakdown voltage V_{CE0} is reached. Once the breakdown has occurred, all the curves tend to merge to the same collector-emitter voltage. The curves merging voltage is denoted as $V_{CE(sus)}$ and it is the minimum voltage necessary to sustain the transistor in breakdown and from Fig. 2.50, the approximate value of $V_{CE(sus)}$ is 115 V.

When operating BJT at high voltage and at high current, another breakdown effect will occur and it is called *second breakdown*. Slight non-uniformities in the current density produce an increased heating of local region. The resistance level of the semiconductor material is decreased because of increased heating in the local region.

The effect results in positive feedback, and the current continues to increase, which further increases the temperature, until the semiconductor material may actually be melt.

The instantaneous power dissipation in a BJT is given by

$$P_Q = V_{CE}I_C + V_{BE}I_B \tag{5.53}$$

where the base current I_B is generally much smaller than the collector current I_C . Therefore, the instantaneous power dissipation is approximated to

$$P_O \cong V_{CE} I_C \tag{2.54}$$

By integrating Eq. (2.54) over one cycle, the average power is given by

$$\overline{P}_{Q} = \frac{1}{T} \int_{O}^{T} V_{CE} I_C dt$$
(2.55)

The average power dissipated in BJT must be kept below a specified maximum value, to ensure that the temperature of the device remains below the maximum value. If the collector current and collector-emitter voltage are DC quantities, then maximum rated power P_T for BJT can be written as

$$P_T = V_{CE} I_C \tag{2.56}$$

The average power limitation P_T is a hyperbola as per the above equation.

The maximum current, voltage and power limitations are illustrated on the I_C versus V_{CE} characteristics as shown in Fig. 2.51(a). Safe Operating Area (SoA) is the region where a transistor can be operated safely and is bounded by I_{Cmax} , $V_{CE(sus)}$, P_T and transistor's second breakdown characteristics. Figure 2.51(b) shows the safe operating area in linear scale and in logarithmic scale. The $I_C - V_{CE}$ operating point may move momentarily outside the safe operating area without damaging the transistor. But this depends on how far and how long the *Q*-point moves outside the area.



Power transistors which are designed to handle large current require large emitter areas to maintain reasonable current densities. These transistors are usually designed with narrow emitter widths to minimize the parasitic base resistance and fabricated as an interdigitated structure as shown in Fig. 2.52. In each emitter leg, small resistors are incorporated which helps to maintain equal current in each *B*-*E* junction.



(b) Fig. 2.52 Interdigitated structure for power transistor

TWO MARK QUESTIONS AND ANSWERS

1. What is semiconductor?

A semiconductor is a material, which is basically an insulator but starts conducting at higher temperatures or by means of doping. The energy band of semiconductor is similar to that of an insulator. Hence, it acts as an insulator at 0K and as a conductor at higher temperatures or by doping.

2. Classify the semiconductor based on chemical composition?

Based on chemical composition, the semiconductors are classified as: (i) Elemental semiconductors and (ii) Compound semiconductors

3. What are degenerate semiconductors?

If doping is very heavy i.e., 10^{23} to 10^{24} atoms/m³, then the conductivity of semiconductors become comparable to metals. Such semiconductors are called degenerate semiconductors. They find applications in high power and high frequency devices.

4. What are called elemental semiconductors?

Semiconductor elements of fourth group, which are doped with pentavalent or trivalent impurities, in order to get *N*-type or *P*-type semiconductors are called elemental semiconductors Example: Ge and Si

2.88 Physics for Electronics Engineering and Information Science

5. What are compound semiconductors?

Semiconductors formed by combining different elements i.e., V & III group or VI & II group elements are called compound semiconductors. Example: GaAs and GaP

6. What is meant by bandgap of a semiconductor?

Bandgap of a semiconductor is the region of energies which are not allowed to occupy the electrons of that material. The impurity atoms can have their energy levels in band gap. It is equal to the energy difference between the minimum energy of conduction band and maximum energy of valence band. It is represented by E_G .

7. Mention the properties of semiconductor.

The properties of semiconductor are as follows:

- (i) Semiconductors are formed by the covalent bonds.
- (ii) They have an empty conduction band.
- (iii) They have almost filled valence band.
- (iv) These materials have a small energy gap.
- (v) They possess crystalline structure.
- (vi) These materials have negative temperature coefficient of resistance.
- (vii) If impurities are added, the electrical conductivity of semiconductor gets increased.
- (viii) If there is an increase in the temperature of semiconductor, its electrical conductivity also increases.

8. Distinguish between conductor and semiconductor.

S. No.	Conductor	Semiconductor
1.	Conductors are those materials which conduct electricity and their conductivity is very high.	Pure semiconductors behave as an insulators at 0 K.
2.	Conductors have low resistivity.	The resistivity of semiconductors lies between conductors and insulators.
3.	In conductors, conduction band and valence band overlap and there is no band gap.	In semi conductors, conduction band and valence band is separated by energy gap E_G .
4.	Examples: Steel, Aluminium and Copper	Examples: Silicon, Germanium and Arsenide

9. Why Silicon is preferred over Germanium?

Silicon is mainly preferred over Germanium because

- (i) In Silicon, the outermost electrons are in 3rd shell whereas in Germanium, they are in 4th shell. So, Silicon is more stable at room temperature than Germanium.
- (ii) Silicon-dioxide (SiO_2) is an insulator which becomes an important part of fabrication during isolation.

10. Write the differences between Silicon (Si) and Germanium (Ge).

Some of the differences between Si and Ge are:

(i) Power handling capability of Si is higher than Ge

- (ii) Easy availability of Si compared to Ge
- (iii) In Si, the operating temperature range is from -60° to 175°C but for Ge, the range is from -60° to 75°C
- (iv) Leakage current is of the order of nA for Si while it is of the order of mA for Ge
- (v) Conductivity of Si is more temperature sensitive than Ge

11. Define intrinsic semiconductor. Give examples.

Semi conducting materials without any impurity atoms or in its pure form are called intrinsic semiconductors. The charge carriers are produced due to breaking of covalent bonds and the concentration of electrons and holes are equal in intrinsic semiconductors. Their energy gap is of the order of 1eV. *Examples:* Si and Ge

12. Write the expression for the electrical conductivity of an intrinsic semiconductor.

The electrical conductivity (σ_i) of an intrinsic semiconductor is given by

$$\sigma_i = n_i q(\mu_n + \mu_p)$$

where n_i is the intrinsic carrier concentration, q is the charge, μ_n and μ_p are the mobilities of electrons and holes, respectively.

13. Why do we prefer Silicon for transistors and GaAs for laser diodes?

Silicon is an indirect bandgap semiconductor for which the life time of the charge carriers is more and the current amplification is also very high. Hence, it is preferable for using it in transistors. GaAs is a direct bandgap semiconductor, in which electrons and holes recombine directly to produce photons and hence, used in laser diodes.

14. Distinguish between direct bandgap and indirect bandgap semiconductors.

S. No.	Direct bandgap	Indirect bandgap
1.	The electron from the conduction band can directly recombine with hole in valence band.	The electron from the conduction band can recombine indirectly with hole in valence band.
2.	Photons are emitted in direct bandgap materials.	Phonons are emitted in indirect bandgap materials.
3.	Lifetime of charge carrier is small.	Lifetime of charge carrier is high.
4.	Examples: GaAs and InP	Examples: Si and Ge

15. Write the difference between intrinsic and extrinsic semiconductor.

S. No.	Intrinsic semiconductor	Extrinsic semiconductor
1.	Intrinsic semiconductor is pure semiconducting material and no impurity atoms are added to it.	Extrinsic semiconductor is prepared by doping a small quantity of impurity atoms of the pure semiconducting materials.
2.	The number of free electrons in the conduction band and the number of holes in the valence band are exactly equal and very small indeed.	The number of free electrons and holes are never equal. There are excess of electrons in <i>N</i> -type semiconductors and excess of holes in <i>P</i> -type semiconductors.

3.	Its electrical conductivity is a function of temperature alone.	Its electrical conductivity depends upon the temperature as well as on the quantity of impurity atoms doped in the structure.
4.	Its electrical conductivity is low.	Its electrical conductivity is high
5.	<i>Examples:</i> Crystalline forms of pure silicon and germanium.	<i>Examples:</i> Silicon "Si" and germanium "Ge" crystals with impurity atoms of As, Sb, P etc. or In B, Al, etc.

16. Define extrinsic semiconductor. Give examples

Semi conducting materials in which the charge carriers originate from impurity atoms added to the pure material are called extrinsic semiconductors. The addition of impurity increases the carrier concentration and hence, the conductivity increases.

Examples: (i) *N*- type semiconductors obtained by adding phosphorous, arsenic or antimony with either Si or Ge and (ii) *P*- type semiconductors obtained by adding Aluminum, Gallium or Indium with either Si or Ge.

17. What are donor impurities?

If pentavalent impurities (P, As and Sb) are doped with pure semiconductors (Si or Ge), it contributes free electrons to the semiconductors. These impurities are called donor impurities.

18. What are acceptor impurities?

If trivalent impurities (B and Ga) are doped with pure semiconductors (Si or Ge), it will accept electrons from the semiconductor. These impurities are called acceptor impurities.

19. What do you mean by doping and doping level?

Doping: The process of adding impurities like Ga, In, and As, etc to a pure or intrinsic semi-conductor is called doping.

Doping level: The maximum extent up to which the impurity can be added is called doping level.

20. What is N-type extrinsic semiconductor?

When pentavalent impurity elements such as phosphorous, arsenic or antimony are added to pure Si or Ge, the resultant semiconductor is called *N*-type semiconductor. Electrons are majority charge carriers in *N*-type semiconductor.

21. What is P-type extrinsic semiconductor?

When trivalent impurity elements such as aluminum, gallium or indium are added to pure Si or Ge, the resultant semiconductor is called *P*-type semiconductor. Holes are majority charge carriers in *P*-type extrinsic semiconductor.

22. What are traps?

Traps are the energy levels situated in the band gap of the semiconductor. This will arises due to impurity atoms. It holds electrons from the conduction band for some times and releases them so as to recombine with holes in the valence band.

23. Explain the concept of holes in semiconductor.

In intrinsic semiconductors, the charge carriers are created due to breaking of covalent bonds. When a covalent bond is broken, an electron escapes to the conduction band leaving behind an empty space in the valence band called a *hole*. These holes will move effectively in the direction opposite to that of the electron.

In *P*-type semiconductors, trivalent impurity elements are added with pure Si or Ge and all the three electrons of impure elements are engaged in forming covalent bonds with three neighboring parent atoms (Si or Ge). Since one more electron is needed to complete its bond, the parent atom supplies one electron, thereby creating a vacant electron site or *hole* on the semiconductor atom.

24. Explain why electrons are majority carriers and holes are minority carriers in an *N*-type semiconductor.

When pentavalent impurities like phosphorus, antimony etc. are added to an intrinsic semiconductor like germanium, it is called *N*-type semiconductor. Here, out of five electrons in phosphorus, four valence electrons makes covalent bonding with the germanium atom and one electron is left free. Similarly, if more number of pentavalent impurities are added, more number of electrons are created and the availability of holes will be very less. Therefore, in *N*-type semiconductors, electrons are majority carriers and holes are minority carriers.

25. Explain why holes are majority carriers and electrons are minority carriers in a *P*-type semiconductor.

When trivalent impurities like boron, indium etc. are added to an intrinsic semiconductor like germanium, it is called *P*-type semiconductor. Here, the three electrons of boron make covalent bonding with the germanium atom and one hole is created. Hence, in *P*-type semiconductors, holes are majority carriers and electrons are minority carriers.

S. No.	N-type semiconductor	P-type semiconductor
1.	<i>N</i> -type semiconductor is obtained by doping an intrinsic semiconductor with pentavalent impurity.	<i>P</i> -type semiconductor is obtained by doping an intrinsic semiconductor with trivalent impurity.
2.	Here, electrons are majority carriers and holes are minority carriers.	Here, holes are majority carriers and electrons are minority carriers.
3.	Donor energy level is close to the conduction band.	Acceptor energy level is close to the valence band.
4.	Electrons are easily donated from the donor energy level to the conduction band, when temperature is increased	Electrons are easily accepted from the valence band to the acceptor energy level, when temperature is increased

26. Compare *N*-type and *P*-type semiconductors.

27. What is the Fermi level in a semiconductor?

Fermi level is the energy level situated in the band gap of the semiconductor.

In Intrinsic semiconductor, the Fermi level is located at the middle of the band gap. It is a reference energy level from which the maximum energy of valence band and minimum energy of conduction band are referred.

In Extrinsic semiconductor, the Fermi level lies in between the acceptor energy level and maximum energy of valence band for *P*-type whereas it lies in between the donor energy level and the minimum energy of conduction band for *N*-type.

28. Give the expression for Fermi energy of an intrinsic semiconductor.

Fermi energy level is the energy level which distinguishes the filled and empty states or it is the maximum energy level up to which the electrons are filled. At 0K, the Fermi energy of an intrinsic semiconductor is given by

$$E_F = \left(\frac{E_V + E_C}{2}\right)$$

The above equation shows that the Fermi energy lies exactly between the lowest energy level of conduction band and highest energy level of valence band.

29. Give the expression for Fermi energy of a *P*-type semiconductor

Fermi energy level is the energy level which distinguishes the filled and empty states or it is the maximum energy level upto which the electrons are filled. At 0K, the Fermi energy of a *P*-type semiconductor is given by

$$E_F = \frac{E_V + E_A}{2}$$

The above equation shows that the Fermi energy lies exactly between the valence band and the acceptor energy level of valence band.

30. Give the expression for Fermi energy of a *N*-type semiconductor

Fermi energy level is the energy level which distinguishes the filled and empty states or it is the maximum energy level up to which the electrons are filled. A 0K, the Fermi energy of an *N*-type semiconductor is given by

$$E_F = \frac{(E_D + E_C)}{2}$$

The above equation shows that the Fermi energy lies exactly between the conduction band and the donor energy level of valence band.

31. Write the expression for carrier concentration in intrinsic semiconductor.

The carrier concentration in intrinsic semiconductor (n_i) is given by

$$n_i = 2 \left(\frac{2\pi kT}{h^2}\right)^{3/2} (m_e^* m_h^*)^{3/4} e^{\left(\frac{-E_G}{2kT}\right)}$$

where m_e^* is the effective mass of electron, m_h^* is the effective mass of hole and E_G is the band gap.

32. Write the expression for carrier concentration in *N*-type semiconductor.

The carrier concentration in *N*-type semiconductor is given by

$$n = (2N_D)^{1/2} \left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/4} e^{\frac{(E_D - E_C)}{2kT}}$$

where m_e^* is the effective mass of electron, N_D is the donor concentration, E_D is the donor energy level and E_C is the energy level in conduction band.

33. Write the expression for carrier concentration in P-type semiconductor

The carrier concentration in *P*-type semiconductor is given by

$$p = (2N_A)^{1/2} \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/4} e^{\left(\frac{E_V - E_A}{2kT}\right)}$$

where m_h^* is the effective mass of electron, N_A is the acceptor concentration, E_A is the acceptor energy level and E_V is the energy level in valence band.

34. Discuss the variation of Fermi level with temperature in the case of *P*-type & *N*-type semiconductor.

In *P*-type semiconductor, the Fermi level shift up and reaches the middle of the band gap by increase in temperature.

In *N*-type semiconductor, the Fermi level shift down and reaches the middle of the band gap by increase in temperature up to 500K.

35. With increase of temperature, the conductivity of semiconductor increases while that of metals decreases. Give reasons.

In semiconductors, more number of charge carriers are created with increase in temperature and hence, the conductivity increases.

In metals or conductors, the concentration of charge carriers remains the same with increase in temperature. But due to increase of thermal energy, electrons make frequent collisions with lattice ions and hence, its resistivity increases and conductivity decreases.

36. What is meant by electron-hole pair?

When the temperature of the semiconductor is increased from 0K, one electron moves from the valence band to the conduction band due to the thermal energy. Hence, a vacant space or hole is created in the valence band. The electron in the conduction and hole created in the valence band is called electron-hole pair.

37. State mass-action law in semiconductors.

Under thermal equilibrium for any semiconductor, the product of the number of holes and the number of electrons is constant and is independent of the amount of donor and acceptor impurity doping. This relation is known as mass-action law and is given by

$$n \cdot p = n_i^2$$

38. Define impurity range in *N*-type semiconductors.

The impurity range is due to the transfer of electrons form donor energy levels to conduction band. Here, the electron concentration in the conduction band steadily increases due to ionization of donor atoms.

2.94 Physics for Electronics Engineering and Information Science

39. Define Exhaustion range in *N*-type semiconductors.

When all the electrons are transferred from donor energy level to conduction band, the electron concentration remains constant over certain temperature and is called exhaustion range.

40. Define intrinsic range in N-type semiconductors.

In intrinsic range, the *N*-type semiconductor practically behaves like an intrinsic semiconductor. Therefore if the temperature is increased, the electron concentration in the conduction band increases rapidly due to the shifting of electrons from valence band to conduction band.

41. What is meant by operating temperature of the semiconductor? Give examples.

Operating temperature of the semiconductor is the maximum temperature up to which the semiconductor behaves as extrinsic semiconductor. Above operating temperature, the extrinsic semiconductor becomes an intrinsic semiconductor. Examples: silicon (doped) transistors can be operated only up to 200°C due to the effect of doping, after which the current amplification will be very less due to its intrinsic behavior.

42. What is meant by carrier transport?

The flow of current is mainly due to the movement of electrons and holes in a semiconductor. The process of movement of these charged particles is called *carrier transport*.

43. Define drift transport and diffusion transport.

The two basic carrier transport mechanisms in a semiconductor material are drift transport and diffusion transport. The movement of charge due to electric fields in a semiconductor is called *drift transport* and the carrier transport through drift results in *drift current*. The flow of charge due to density gradients in a semiconductor is called *diffusion transport* and the carrier transport by diffusion results in *diffusion current*.

44. Define diffusion current and drift current.

Drift current: The drift current is defined as the flow of electric current due to the motion of the charge carriers under the influence of an external electric field. The drift current is given by

$$I_n = Aqn\mu_0 E$$
 (due to free electrons)
 $I_p = Aqp\mu_0 E$ (due to holes)

Diffusion current: A concentration gradient exists if the number of either electrons or holes is greater in one region of a semiconductor, compared to the rest of the region. The charge carriers tend to move from the region of higher concentration to that of lower counteraction of the same type of charge carriers. There, the movement of charge carriers takes place resulting in a current called diffusion current. The diffusion current is given by

$$I_n = A \cdot qD_n \frac{dn}{dx} \qquad (\text{due to free electrons})$$
$$I_p = -A \cdot qD_p \frac{dp}{dx} \qquad (\text{due to holes})$$

45. State Einstein's relationship for semiconductor.

There exists a definite relationship between the mobility and diffusion coefficient of a particular type of charge carrier in the same semiconductor. The higher the value of mobility of a charge carrier, the greater will be its tendency to diffuse. The equation which relates the mobility μ and the diffusion coefficient *D* is known as the *Einstein's Relationship*. The Einstein's relationship is expressed as

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = \frac{kT}{q} = V_T$$

The importance of Einstein's relationship is that it can be used to determine D_p (or D_n), if the mobility of holes (or electrons) is measured experimentally.

46. Define electron volt.

1 eV is the amount of energy acquired by an electron when it is accelerated through a potential difference of one volt (1 V). The value of 1 eV is equal to 1.602×10^{-19} joule.

47. What is depletion region in *PN* junction?

The region around the junction from which the mobile charge carriers i.e., electrons and holes are depleted is called as depletion region. Since this region has immobile ions, which are electrically charged, the depletion region is also known as space charge region.

48. Define barrier potential.

With no applied voltage across the *PN* junction diode, a barrier is set-up across the junction which prevents movement of charge carriers i.e., electrons and holes. This leads to induced electric field across the depletion layer, and an electrostatic potential difference is established between the *P*-and *N*-regions. Such potential is called the barrier potential, junction barrier, diffusion potential, or contact potential, V_{o} .

49. What is forward bias and reverse bias in a PN junction?

When positive terminal of the external supply is connected to P region and negative terminal to N region, then the PN junction is said to be forward biased. Under forward biased condition, the PN region offers a very low resistance and a large amount of current flows through it.

When positive terminal of the external supply is connected to N region and negative terminal to P region, then the PN junction is said to be in reverse biased. Under reverse biased condition, the PN region offers a very high resistance and a small amount of current flows through it.

50. Define avalanche breakdown.

Under reverse bias condition, when the reverse voltage approaches a particular value called breakdown voltage, the reverse current suddenly increases because the velocity of thermally generated minority carriers increases and they cross the depletion region and acquire sufficient kinetic energy from the applied potential to produce new charge carriers by removing valence electrons from their bonds. These new carriers will in turn collide with other atoms and will increases the number of electron-holes available for conduction. Because of the cumulative increase in carrier density after each collision, this process is called *avalanche breakdown*.

2.96 Physics for Electronics Engineering and Information Science

51. What is Zener breakdown?

When the P and N regions are heavily doped, direct rupture of covalent bonds takes place because of the strong electric field at the junction of PN diode. The new electron-hole pair, so created, increases the reverse current in a reverse biased PN diode. As a result of heavy doping of P and N regions, the depletion width becomes very small and the field across the depletion region becomes very high and it is due to ruptures of the covalent bond. This breakdown is known as *Zener breakdown*.

52. What is Hall effect and Hall voltage?

When a piece of conductor (metal or semiconductor) carrying current is placed in a transverse magnetic field, an electric field is produced inside the conductor in a direction normal to both currents and magnetic field. This phenomenon is known as Hall effect and the voltage generated is called as Hall voltage.

53. List the applications of Hall effect.

Hall effect is used

- (i) To determine the type of semiconductor whether it is N-type or P-type
- (ii) To measure carrier concentration
- (iii) To determine mobility of charge carriers
- (iv) To measure magnetic flux density using a semiconductor sample of known Hall- Coefficient
- (v) To determine electrical conductivity of the semiconductors
- (vi) To determine whether the solid is metal, insulator or semiconductor

54. Define Hall voltage.

The voltage developed between plates of a current carrying conductor, when placed in transverse magnetic field is called as hall voltage.

55. Write an expression for hall coefficient.

Hall coefficient, $R_H = -\frac{1}{nq}$, for *N*-type semiconductor

and

$$R_H = \frac{1}{nq}$$
, for *P*-type semiconductor

where n is the carrier concentration of the respective charge carriers of P-type and N-type semiconductor.

56. How to determine the type of semiconductor based on Hall effect?

If the Hall coefficient R_H is negative, then the semiconductor is an *N*-type semiconductor. If the Hall coefficient R_H is positive, then the semiconductor is a *P*-type semiconductor.

57. Write an expression for Hall coefficient.

A rectangular slab of a given material having a thickness t and width b is taken. The slab is placed between two pole pieces of an electromagnet such that the magnetic field B is perpendicular to current I. The Hall voltage V_H is then measured by placing two probes at the two side faces of the slab. If the magnetic flux density is B and the Hall voltage is V_H , then the Hall coefficient R_H is given by

$$R_H = \frac{V_H.t}{IB}$$

58. What is a metal-semiconductor contact?

A metal-semiconductor contact is a junction contact of a metal with a piece of semiconductor. The active junction is the interface between the metal, which acts as an anode, and the semiconductor. The other interface between the semiconductor and the metal, which acts as a cathode, is an Ohmic contact and there is no potential drop at this junction.

59. What is tunneling phenomenon?

According to the classical laws of physics, a charged particle should possess energy at least equal to the energy barrier in order to cross an energy barrier. Hence, the particle will cross the energy barrier if its energy is greater than the barrier and cannot cross the barrier if its energy is less than the energy barrier. But quantum mechanically, there exists non-zero probability that the particle with energy less than the energy barrier as if it tunnels across the barrier. This is called *tunneling effect*.

60. List out the applications of tunnel diode.

Tunnel diode is used:

- (i) as an ultra-high speed switch with switching speed of the order of ns or ps,
- (ii) as a logic memory storage device,
- (iii) as a microwave oscillators,
- (iv) in relaxation oscillator circuit,
- (v) as an amplifier,
- (vi) microwave and RF power monitors, and
- (vii) high-frequency triggers.

61. Mention some advantages and disadvantages of Tunnel Diode.

Advantages of Tunnel Diode are:

- (i) Low noise
- (ii) Ease of operation
- (iii) High speed and
- (iv) Low power

Disadvantages of Tunnel Diode are:

- (i) Voltage range over which it can be operated is less than 1 V.
- (ii) Being a two terminal device, there is no isolation between the input and output circuit.

62. What are the differences between a tunnel diode and an ordinary PN junction diode?

Unlike an ordinary PN junction diode, the tunnel diode is heavily doped which results in a very thin depletion region in terms of few Angstrom. The high electric field between the P and N regions causes tunneling of charge carriers resulting in a current flow with a very small bias voltage. Further, the tunnel diode exhibits negative resistance characteristics wherein the conduction current decreases with increase in the applied voltage.

2.98 Physics for Electronics Engineering and Information Science

63. Compare Schottky diode and conventional diode.

- (i) The current in a conventional PN junction diode is controlled by the diffusion of minority carriers whereas the current in the Schottky diode results from the flow of majority carriers over the potential barrier at the metal-semiconductor junction.
- (ii) The reverse saturation current for a Schottky diode is larger than that of a PN junction diode.
- (iii) The Schottky diode has a smaller turn-on voltage and shorter switching time than the PN junction diode.

64. What is NMOS capacitor?

The MOS capacitor consists of a Metal-Oxide-Semiconductor structure in which, the semiconductor P-type substrate which is the bottom electrode formed a N^+ layer with a thin oxide layer (SiO₂) and a top metal contact, referred to as the gate. A second metal layer forms an ohmic contact to the back of the semiconductor and is called the bulk contact. This is known as N-type MOS or NMOS capacitor since the inversion layer contains electrons.

65. Give the three different operating modes in the MOS capacitor.

The three different operating modes in the MOS capacitor are (i) accumulation, (ii) depletion and (iii) inversion.

66. Define power transistors.

Power transistors are those which can handle a large amount of current and also dissipates large amount of power across collector-base junction.

67. What are the types of power transistors?

Power transistors are of three types. They are:

- (i) Bipolar Junction Transistor (BJT)
- (ii) Metal Oxide Semiconductor Field Effect Transistor (MOSFET)
- (iii) Insulated Gate Bipolar Transistor (IGBT)

REVIEW QUESTIONS

- 1. What are the three commonly used semiconductors?
- 2. What are the band gap energies of germanium, silicon and gallium arsenide?
- 3. Explain how electrons are emitted from metals.
- 4. Distinguish between elemental and compound semiconductor.
- 5. What is meant by intrinsic semiconductor?
- 6. With covalent bond and energy band diagram, explain the intrinsic semiconductor.
- 7. Describe briefly about direct and indirect bandgap semiconductor.
- 8. Derive an expression for density of electrons in conduction band in an intrinsic semiconductor.
- 9. Derive an expression for density of holes in valence band in an intrinsic semiconductor. What is the effect of temperature on Fermi level in an intrinsic semiconductor?
- 10. Starting with the conductivity of charge carrier in an intrinsic semiconductor, describe how you will determine the band gap of an intrinsic semiconductor.

- 11. Explain the variation of Fermi level with temperature in intrinsic semiconductor.
- 12. Derive an expression for carrier concentration in an intrinsic semiconductor.
- 13. Write the expression for electron and hole concentration in an intrinsic semiconductor and hence derive the

expression: $E_F = \frac{3kT}{4} \ln\left(\frac{m_h^*}{m_e^*}\right) + \left(\frac{E_V + E_C}{2}\right)$ for Fermi level in the intrinsic semiconductor. Assume the symbols to

have their usual meanings.

- 14. Draw a neat sketch of a band diagram of intrinsic semiconductor at room temperature and show that the Fermi level in an intrinsic semiconductor lies in the middle of the energy gap.
- 15. Explain in brief the concept of Fermi level. Show diagrammatically the Fermi level in metals, intrinsic semiconductor and insulators at oK and at higher temperatures.
- 16. Using the expressions of electron concentration and hole concentration for an intrinsic semiconductor, show that the intrinsic carrier density is independent of Fermi level.
- 17. Obtain an equation for the conductivity of an intrinsic semiconductor in terms of carrier concentration and carrier mobility.
- 18. Derive an expression for electrical conductivity of an intrinsic semiconductor and also determine the band gap.
- 19. If effective mass of an electron is equal to twice the effective mass of hole, determine the position of the Fermi level in an intrinsic semiconductor from the centre of forbidden gap at room temperature.

(Answer: The Fermi level is 0.014 eV below the centre of forbidden gap)

- 20. Explain what a hole is. How do they move in intrinsic semiconductor?
- 21. The mobility of electrons and holes in a sample of intrinsic germanium at room temperature are 0.36 m² / V·s and 0.17 m² / V·s, respectively. If the electron and hole densities are each equal to 2.5 × 10¹⁹/m³, calculate the conductivity. (Answer: 2.12 S/m)
- 22. The mobility of free electrons and holes in pure silicon are 0.13 and 0.05 m² / V · s and the corresponding values for pure germanium are 0.38 and 0.18 m² / V · s respectively. Determine the values of intrinsic conductivity for both silicon and germanium. Given that $n_i = 2.5 \times 10^{19}$ /m³ for germanium and $n_i = 1.5 \times 10^{16}$ /m³ for silicon at room temperature. (Answer: 0.43 S/m; 2.24 S/m)
- 23. The resistivity of intrinsic silicon at 270°C is 300 $\Omega \cdot m$. Calculate the intrinsic carrier density. Assume; $\mu_n = 0.17$ m²/V·s and $\mu_n = 0.035$ m²/V·s. (Answer: 1.0 × 10¹⁶ /m³)
- 24. Determine the fraction of electrons in conduction band in silicon at 27°C and 227°C. Given: $E_G = 11 \text{ eV}$ and $k = 1.38 \times 10^{-23} \text{ J/K}$. (Answer: 5.7×10^{-10} ; 2.8×10^{-6})
- 25. Calculate the temperature at which Silicon (E_G = 1.14 eV) will have the same fraction of electrons in the conduction band as Germanium (E_G = 0.72 eV) has at 300 K.
 (Answer: 475 K)
- 26. Find the fraction of electrons in the valence band of intrinsic geranium which can be thermally excited across the forbidden energy gap of 0.7 eV into the conduction band at (i) 50 K (ii) 300 K and (iii) 1000 K.

(Answer: 5.1×10^{-36} ; 1.3×10^{-6} ; 0.017)

- 27. What fraction of the conductivity of intrinsic silicon at room temperature is due to (i) electrons and (ii) holes? The electrons and the hole mobility's are, $\mu_n = 0.135 \text{ m}^2 / \text{V} \cdot \text{s}$ and $\mu_p = 0.048 \text{ m}^2 / \text{V} \cdot \text{s}$. (Answer: 73.8% due to electrons and 26.2% due to holes)
- 28. Compute the current produced in a Ge plate of 1 cm² area, 0.03 mm thickness across its faces. Assume a free electron concentration of 2×10^{19} m³, the electron and hole mobility's being 0.39 m² / V s and 0.19 m² / V s respectively. (Answer: 12.4 A)
- 29. The energy gap in silicon crystal is 1.12 eV. Its electron and hole mobilities at room temperature are 0.48 m²/V·s and 0.013 m²/V·s respectively. Find its conductivity. (Answer: $1.29 \times 10^{-3} (\Omega \cdot m)^{-1}$)

2.100 Physics for Electronics Engineering and Information Science

- 30. Mobilities of electron and holes in a sample of intrinsic germanium at room temperature are 3600 cm² / V · s and 1700 cm² / V · s respectively. If the electron and hole densities are each equal to 2.5×10^{13} / cm³, calculate the conductivity. (Answer: $2.12 (\Omega \cdot m)^{-1}$)
- 31. The hole concentration in silicon varies linearly from x = 0 to x = 0.01 cm² / s. The hole diffusion current density is 20 A/cm² and the hole concentration at x = 0 is $p = 4 \times 10^{17}$ cm⁻³, What is the value of the hole concentration at x = 0.01 cm² (Answer : 2.75×10^{17} cm⁻³)
- 32. Why are holes not generated in metals? Can the electrical conduction of semiconductors be improved to the extent possible in metals? Discuss.
- 33. What is meant by doping in a semiconductor?
- 34. Discuss the following with respect to semiconductor: (i) doping (ii) dopant (iii) donor and (iv) acceptor.
- 35. Explain "majority and minority carriers" in a semiconductor.
- 36. Explain the differences between intrinsic and extrinsic semiconductors.
- 37. What is meant by *N*-type impurity in a semiconductor?
- 38. What is meant by P-type impurity in a semiconductor?
- 39. Differentiate between *N*-type and *P*-type semiconductors.
- 40. What do you understand by intrinsic and extrinsic semi-conductors?
- 41. With covalent bond and energy band diagram, explain *N*-type and *P*-type extrinsic semiconductor.
- 42. Draw energy band diagrams for *N*-type semiconductor at 0° K and at T° K
- 43. Derive the expression for the number of electrons per unit volume in conduction band of *N*-type extrinsic semiconductor.
- 44. Discuss the variation of carrier concentration with temperature.
- 45. Describe the variation of the Fermi level with temperature in case of *N*-type semiconductor.
- 46. Derive the expression for the number of holes per unit volume in valence band of *P*-type extrinsic semiconductor.
- 47. How does the Fermi level change with temperature in extrinsic semiconductors?
- 48. Explain with sketch the variation of Fermi level and carrier concentration with temperature in the case of *P* and *N* type semiconductors for high and low doping levels.
- 49. Discuss the effect of increasing amounts of dopants on the Fermi level in extrinsic semiconductors.
- 50. How does carrier concentration vary with temperature in an extrinsic semiconductor?
- 51. An *N*-type extrinsic semiconductor is in equilibrium at room temperature. What is the net charge on the impurity atom? Discuss.
- 52. Explain the effect of impurity concentration on the Fermi level in an extrinsic semiconductor.
- 53. How does the Fermi level change with temperature in extrinsic semiconductors? Discuss the effect of increasing amounts of dopants in extrinsic semiconductors.
- 54. Explain the variation of Fermi level position with temperature and donor concentration in *N*-type and *P*-type semiconductor.
- 55. A silicon wafer is doped with 10^{21} phosphorus atoms /m³. Calculate (i) the majority carrier concentration, (ii) the minority concentration and (iii) the electrical resistivity of the doped silicon at room temperature. Assume complete ionization of the dopant atoms; $n_i = 1.5 \times 10^{-16}$ /m³, $\mu_n = 0.135$ m² / V·s and $\mu_p = 0.048$ m² / V·s. (Answer: 10^{21} /m³; 2.25×10^{11} /m³; $0.046 \Omega \cdot m$)
- 56. A sample of intrinsic silicon at room temperature has a carrier concentration of 1.5×10^{16} / m³. If a donor impurity is added to the extent of 1 donor atom per 10^8 atoms/m³, determine the resistivity of the material. Given: $\mu_n = 0.135 \text{ m}^2/\text{V} \cdot \text{s}$ and $\mu_p = 0.048 \text{ m}^2/\text{V} \cdot \text{s}$ (Answer: $0.09 \Omega \cdot \text{m}$)

- 57. A current density of 10^3 A/m^2 flows through an *N*-type germanium crystal which has resistivity of $0.05 \Omega \cdot \text{m}$ Calculate the time taken for electrons in material to drift a 5×10^4 m distance. The mobility of electrons is $0.38 \text{ m}^2/\text{V} \cdot \text{s}$ (Answer: $2.6 \times 10^{-6} \text{ S}$)
- 58. A sample of Germanium is doped to the extent of 10^{20} donor atoms/m³ and 7×10^{19} acceptor atoms /m³. At the room temperature of the sample, the resistivity of intrinsic germanium is 0.6 $\Omega \cdot m$ If the applied electric field is 200 V/m, find the total conduction current density. Assume $\mu_n = 0.38 \text{ m}^2/\text{V} \cdot \text{s}$ and $\mu_p = 0.18 \text{ m}^2/\text{V} \cdot \text{s}$ (Answer: 431 A/m²)
- 59. In an *N*-type semiconductor, the Fermi level lies 0.4 eV below the conduction band at 300 K. Determine the new position of the Fermi level if (i) the temperature is increased to 400 K and (ii) the concentration of donor atoms is increased by a factor of 6. Assume *kT* = 0.03 eV.

(Answer: (i) 0.533 eV and (ii) 0.3463 eV below the conduction band)

- 60. Calculate the conductivity of a pure silicon at room temperature of 300 K. Given that $n_i = 1.5 \times 10^{16}$ /m³, $m_n = 0.13 \text{ m}^2/\text{V-s}$, $m_p = 0.05 \text{ m}^2/\text{V} \cdot \text{s}$ and $q = 1.602 \times 10^{-19}$ C. Now the silicon is doped 2 in 10⁸ of a donor impurity. Calculate its conductivity if there are 5×10^{28} silicon atoms/m³. By what factor has the conductivity increase? (Answer: 4.32×10^{-4} S/m; 20.8 S/m; 48,000)
- 61. In a *P*-type semiconductor, the Fermi level lies 0.4 eV above the valence band at 300 K. Determine the new position of the Fermi level (a) at 450 K and (b) if the concentration of acceptor atoms is multiplied by a factor of 2. Assume *kT* = 0.03 eV.
 (Answer: (a) 0.6 eV and (b) 0.38 eV above the valence band)
- 62. Given that $n_i = 1.4 \times 10^{16}$ / m³, $\mu_e = 0.145$ m² / V · s and $\mu_h = 0.05$ m² / V · s. Compute the conductivity of a silicon semiconductor which is doped with acceptor impurity to a density of 10^{22} atoms/m³. (Answer: 80 S/m)
- 63. The conductivity of pure silicon at room temperature is 5×10^{-4} S/m. How many aluminium atoms per m³ are required so that a saturation conductivity of 200 S/m could be achieved in silicon using aluminium as an impurity? Given that the mobility of holes in Silicon is $0.05 \text{ m}^2/\text{V} \cdot \text{s}$ and the mobility of electrons is $0.13 \text{ m}^2/\text{V} \cdot \text{s}$. (Answer: $2.5 \times 10^{22}/\text{m}^3$)
- 64. Write short notes on velocity-electric field relations in semiconductors.
- 65. Explain the drift and diffusion transport for a semiconductor.
- 66. Differentiate between drift and diffusion currents.
- 67. State and explain Mass-action law.
- 68. What is Einstein's relationship in a PN junction?
- 69. What is a PN junction? How is it formed?
- 70. Explain the formation of depletion region in a PN junction.
- 71. Draw the energy band diagram of a PN junction and explain the working of a diode.
- 72. Sketch the conduction and valence bands before and after diffusion of carriers in a PN junction.
- 73. Explain how a barrier potential is developed at the PN junction.
- 74. Explain V–I characteristics of a PN junction diode.
- 75. Describe the action of *PN* junction diode under forward bias and reverse bias.
- 76. Explain Avalanche breakdown and Zener breakdown in PN junctions.
- 77. What is Hall effect?
- 78. Define Hall coefficient and Hall voltage.
- 79. Explain Hall effect. How can Hall effect be used to determine some of the properties of a semiconductor?
- 80. Describe an experimental setup to determine the Hall coefficient.
- 81. Obtain an expression for Hall coefficient. Does RH depend on the doping concentration?
- 82. How does Hall effect show whether holes or electrons predominate in a semiconductor?
- 83. Describe the applications of Hall effect.

2.102 Physics for Electronics Engineering and Information Science

- 84. Derive an expression for charge density in terms of Hall voltage and further explain how mobility of the charge carriers can be evaluated by knowing the conductivity.
- 85. A sample of *N*-type semiconductor has a Hall coefficient of 150 cm³ / coulomb. If its resistivity is 0.15 $\Omega \cdot m$, estimate the electron mobility in the sample. (Answer: 1000 cm²/V·s)
- 86. The conductivity of a pure silicon bar is 5 × 10⁻⁴ (Ω · m)⁻¹. The magnetic flux density is 0.1 Wb/m². and the thickness of the bar in the direction of the magnetic field is 3 mm. The measured values of Hall voltage and current are 50 mV and 10 µA, respectively. Find the hole mobility.

(Answer: $0.075 \text{ m}^2/\text{V} \cdot \text{s}$)

- 87. A semiconducting crystal 12 mm long, 5 mm wide and 1 mm thick has a magnetic flux density of 0.5 Wb/m² applied from front to back perpendicular to the largest faces. When a current of 20 mA flows lengthwise through the specimen, the voltage measured across its width is found to be 37. What is the Hall coefficient of this semiconductor? $(Answer: 3.7 \times 10^{-6} m^3 / C)$
- 88. An N-type Ge sample has a donor density of 10²¹ m³. It is arranged in Hall effect experiment having magnetic field of 0.5 Tesla and current density 500 A/m². Find the hall voltage, if the sample is 3 mm wide. (Answer: 4.7 mV)
- 89. An *N*-type semiconductor has Hall coefficient $4.16 \times 10^{-4} \text{ m}^3 / \text{C}$. The conductivity is $108 (\Omega \cdot \text{m})^{-1}$. Calculate its charge carrier density n_c and electron mobility at room temperature. (Answer: $4.5 \times 10^{-4} \text{ m}^2 / \text{V} \cdot \text{s}$)
- 90. Draw the structure of a metal–semiconductor junction and explain the energy band structure before and after contact.
- 91. Derive an expression for built-in potential in a metal and *N*-type semiconductor junction in terms of barrier height and carrier concentration.
- 92. Derive an expression for built-in potential in a metal and *P*-type semiconductor junction in terms of barrier height and carrier concentration.
- 93. Describe ohmic contact with its energy band diagram.
- 94. What is tunneling?
- 95. From the energy band diagram, explain the V–I characteristic of a tunnel diode.
- 96. Draw the equivalent circuit of a tunnel diode and explain it.
- 97. List out the applications of tunnel diode and mention its advantages and disadvantages.
- 98. What is a Schottky diode? Explain the flow of carriers across its junction during forward and reverse biased conditions with energy band diagrams.
- 99. Briefly explain how the construction of a Schottky diode favours its use in the high frequency region.
- 100. Draw the C-V characteristics of the MOS capacitor.
- 101. Explain the different operating modes in the MOS capacitor.
- 102. Define Safe Operating Area (SOA) in power transistors.
- 103. Describe briefly about power transistor with its V-I characteristics.

Magnetic and Dielectric Properties of Materials

3.1 INTRODUCTION

Magnetic materials play a prominent role in modern technology. They are widely used in electronics and computer industries. Many electrical devices such as inductors, transformers, rotating machines and ferrite antennas are based on utilizing the magnetic properties of materials. The soft and hard magnetic properties of materials are used in computer data storage. Magnetic materials are those substances when placed in an external magnetic field also become source of a magnetic field. The magnetic materials are broadly classified into three groups, namely diamagnetic, paramagnetic and ferromagnetic materials.

Dielectric materials are used for electrical insulations in cables, wires etc. in the form of coating around them. The dielectric materials are also used in electronic equipment such as rectifiers, transducers, and amplifiers. The solid dielectric materials with high permittivity are used in capacitors to store electric charge at a given voltage. Liquid dielectrics like mineral oil, silicone oil, or organic ester that have the ability to insulate the coils of a transformer both electrically and thermally are also employed as electrical insulators. The gaseous dielectrics such as air, nitrogen and sulfur hexafluoride are commonly used.

Some dielectric materials like Barium Titanate have very high dielectric constants due to their high value of spontaneous dipole moment and are termed as ferroelectric materials. Also, another property of dielectric materials namely, the dielectric loss which is dissipated in the form of heat is applied to food processing technique in microwave cooking. This chapter deals with the properties of magnetic and dielectric materials in detail.

3.2 MAGNETISM IN MATERIALS

Origin of Magnetization

The magnetic properties of solids arise due to the presence of atomic dipoles which give rise to magnetic dipole moments. In general, the magnetic dipole moment of the atom arises from three sources as discussed below:

The Orbital Motion of Electrons

Any material is made up of atoms in which the electrons revolve in orbits around the nucleus as shown in Fig. 3.1.

3.2 Physics for Electronics Engineering and Information Science

The phenomenon of orbiting of electrons produces an internal magnetic field. This field is similar to the magnetic field produced by a current loop. The equivalent current loop exhibits a magnetic moment and the current loop is regarded as a magnetic dipole. The direction of magnetic moment is normal to the plane of the loop in accordance with the right hand rule. The total orbital magnetic moment of an atom is the sum of orbital magnetic moments of individual electrons.



Fig. 3.1 Orbital motion of an electron

Electron Spin

Each electron is spinning about an axis through itself either clockwise or anticlockwise as shown in Fig. 3.2 and this spin also gives rise to a magnetic dipole moment called spin magnetic moment.



Fig. 3.2 Clockwise and anticlockwise motion of electrons

Nuclear Spin

In addition to electronic contribution, nuclear spin also contributes to magnetic moment of atoms. The magnetic moment of the nucleus is about 1/2000th of the magnetic moment of the electron. Hence, the magnetic moment due to nuclear spin is neglected. The magnetization in any material is associated with atomic current loops generated by orbital motion of electrons around the nucleus and spin motion of electrons.

3.2.1 Magnetic Dipole Moment – Bohr Magneton

Bohr magneton is the elementary electron magnetic moment because no electron can have a magnetic moment below it. It is represented by

$$\mu_B = \frac{e\hbar}{2m_e} = 9.27 \times 10^{-24} \,\mathrm{A} \cdot \mathrm{m}^2$$

An orbiting electron in an atom behaves like a current loop and has a magnetic dipole moment associated with it and is called the orbital magnetic moment ($\overline{\mu}_{orb}$), as illustrated in Fig. 3.3. If ω is



Fig. 3.3 An orbiting electron equivalent to a magnetic dipole moment $\bar{\mu}_{\rm orb}$

the angular frequency of the electron, then the current I due to the orbiting electron is

I = Charge flowing per unit time
$$= -\frac{e}{\text{period}} = -\frac{e\omega}{2\pi}$$

If r is the radius of the orbit, then the magnetic dipole moment is

$$\overline{\mu}_{\rm orb} = I(\pi r^2) = -\frac{e\omega r^2}{2}$$

But the velocity v of the electron is equal to $r\omega$ and its orbital angular momentum is

$$L = (m_e v)r = m_e \omega r^2$$

Substituting the above equation in μ_{orb} , we get

$$\overline{\mu}_{\rm orb} = -\frac{e}{2m_e}L\tag{3.1}$$

It is seen that the magnetic moment is proportional to the orbital angular momentum through a factor that has the charge to mass ratio of the electron. The numerical factor $e/2m_e$, relating the angular momentum to the magnetic moment, is called the *gyromagnetic ratio*. The negative sign in Eqn. (3.1) indicates that, μ_{orb} is in the opposite direction to L as shown in Fig. 3.3 which is due to the negative charge of the electron.

The electron also has an intrinsic angular momentum or spin, S. The spin of the electron has a spin magnetic moment, denoted by μ_{spin} , but the relationship between μ_{spin} and S is not the same as that in Eqn. (3.1), the gyromagnetic ratio is 2 times greater and is given by

$$\mu_{\rm spin} = -\frac{e}{m_e}S$$

Quantum mechanics requires the spin angular momentum to be space quantized. Hence,

$$S = s\hbar$$

where s = 1/2 is the spin quantum number. Thus, the spin magnetic moment is

$$\mu_{\rm spin} = -\frac{e}{m_e}S = -\frac{e}{m_e}s\hbar = -\frac{e\hbar}{2m_e}$$

The quantity $\mu_B = \frac{e\hbar}{2m_e}$ is called the Bohr magneton and has the value $9.27 \times 10^{-24} \,\text{A} \cdot \text{m}^2$. Thus, the spin of a single electron has a magnetic moment of one Bohr magneton.

The nuclear magneton which arises from the spin of the nucleus is given by

$$\mu_p = \frac{e\hbar}{2m_p}$$
 (i.e, in analogy with Bohr magneton)

where, m_p is the mass of proton , which is 10^3 times the mass of an electron. Substituting their respective values, we get

$$\mu_n = 5.05 \times 10^{-27} \,\mathrm{A \cdot m^2}$$

or, $\mu_p = 10^{-3} \mu_B$ and hence for all practical purposes, it may be ignored.

In general, the orbital magnetic moments also cancel each other due to their random orientations and the spin magnetic moment decides whether the material contains permanent dipoles or not and the types of material.

3.3 MAGNETIC FIELD AND INDUCTION

Magnetic Field, \overline{H}

Magnetic field is produced by permanent magnets like horse shoe magnets and temporarily by electromagnets or superconductor magnets. Since the magnetic field is a vector quantity, it is denoted by \overline{H} . The unit of \overline{H} is ampere-turns per metre (A/m) in SI system.

Magnetic Induction or Magnetic Flux Density \overline{B}

A magnetic field is schematically represented by lines of magnetic induction. The lines of induction are collectively called flux. The number of field or flux lines (ϕ) passing through unit area of cross-section A is called the magnetic flux density \overline{B} (vector quantity). It is given by

$$\overline{B} = \frac{\text{Magnetic flux}}{\text{area}} = \frac{\phi}{A}$$

The quantity \overline{B} is measured in Weber per square metre (Wb/m²) or tesla (T).

3.4 MAGNETIZATION

The dipole moments are randomly oriented in a material as shown in Fig. 3.4(a). The magnetic dipole moment is zero without an external magnetic field applied to the material. The magnetic behaviour of a material is characterized by the interaction of the magnetic dipole moments of its atoms with the external magnetic field. However, when such a material is kept in an external magnetic field shown in Fig. 3.4 (b), the random orientation of dipole moments disappears and they tend to align themselves in the direction of the applied field. As a result, the net magnetic moment is not zero. Thus, the material is magnetized.



Fig. 3.4 Magnetic dipole moment in a differential volume Δv (a) in the absence of external field and (b) in the presence of external field

Magnetization \overline{M}

It is defined as the magnetic moment per unit volume developed inside a solid. In SI system, \overline{M} is measured in amperes per metre (A/m). Since the magnetization is induced by the field, it is assumed that \overline{M} is proportional to \overline{H} and it is expressed as

$$\overline{M} \propto \overline{H}$$

 $\overline{M} = \chi_m \overline{H} \tag{3.2}$

where χ_m is the proportionality constant and is called magnetic susceptibility.

3.5 MAGNETIC PERMEABILITY AND MAGNETIC SUSCEPTIBILITY

Relationship between \overline{B} and \overline{H}

When a material is kept in a magnetic field, two types of induction arise, one due to the magnetizing field, \overline{H} and the other as a consequence of the magnetization, \overline{M} of the material itself. Therefore, the magnetic induction, \overline{B} , produced inside the material is given by

$$\overline{B} = \mu_0 (\overline{H} + \overline{M}) \tag{3.3}$$

where μ_0 is known as the permeability of the free space. It is equal to $4\pi \times 10^{-7}$ Henry per metre (H/m).

From Eqns.(3.2) and (3.3), we get

$$\overline{B} = \mu_0 (1 + \chi_m) \overline{H} \tag{3.4}$$

or

$$\overline{B} = \mu \overline{H} \tag{3.5}$$

where μ is called the absolute permeability of the medium and in an isotropic medium, it is a scalar quantity.

In case of free space, $\overline{M} = 0$ and Eqn. (3.3) becomes

$$\overline{B} = \mu_0 \overline{H}$$

Absolute Permeability, μ

When a magnetic material is placed in a magnetic field, the magnetic field lines are redistributed and tend to pass more (or less in some cases) through the material. The absolute permeability of the material is a measure of the degree of which the field lines penetrate or permeate the material. It is defined as the ratio of the magnetic induction, \overline{B} , in the medium to the magnetizing field, \overline{H} . Thus,

$$\mu = \frac{\overline{B}}{\overline{H}}$$

The unit of absolute permeability is Henry per metre (H/m).

Relative Permeability μ_r

It is defined as the ratio of the absolute permeability of the material to the permeability of free space and it is given by

$$\mu_r = \frac{\mu}{\mu_0} \tag{3.6}$$

where μ_r is the relative permeability and has no unit. For air or vacuum, $\mu_r = 1$.

Magnetic Susceptibility, χ_m

The magnetic susceptibility is a constant that indicates the degree of magnetization of a material in response to applied magnetic field. It is a measure of the ease with which the material can be magnetized. It is defined as the magnetization produced in the material per unit applied magnetic field. It is denoted by χ_m and is given by

$$\chi_m = \frac{\overline{M}}{\overline{H}}$$

In isotropic media, \overline{M} and \overline{H} point in the same direction and χ_m is a scalar quantity. Materials having high susceptibility are easily magnetized.

Relation between μ_r and χ_m

Comparing Eqns. (3.4) and (3.5), we find that

$$\mu = \mu_0 (1 + \chi_m)$$

Using Eqn. (3.6) in the above equation, we obtain

$$\mu_r = (1 + \chi_m) \tag{3.7}$$

Equation (3.7) relates the relative permeability and the susceptibility of the material.

EXAMPLE 3.1

The magnetic field intensity H = 1200 A/m is in a material when B = 2 Wb/m². When H is reduced to 400 A/m, B = 1.4 Wb/m². Calculate the change in magnetization M.

Solution

Case (i)

Given H = 1200 A/m and B = 1.4 Wb/m².

Since $B = \mu H$, the permeability is

$$\mu = \frac{B}{H} = \frac{2}{1200} = \frac{1}{600}$$

Relative permeability, $\mu_r = \frac{\mu}{\mu_0} = \frac{1}{600} \times \frac{1}{4\pi \times 10^{-7}} = 1326.96$

Magnetic susceptibility, $\chi_m = \mu_r - 1 = 1325.96$

Magnetization, $M_1 = \chi_m H = 1591.152 \times 10^3 \, \text{A/m}$

Case (ii)

Given H = 400 A/m and B = 1.4 Wb/m².

Permeability,

$$\mu = \frac{B}{H} = \frac{1.4}{400}$$

Relative permeability, $\mu_r = \frac{\mu}{\mu_0} = \frac{1.4}{400} \times \frac{1}{4\pi \times 10^{-7}} = 2786.6$
Magnetic susceptibility, $\chi_m = \mu_r - 1 = 2785.6$

Magnetization, $M_2 = \chi_m H = 1114.24 \times 10^3 \text{ A/m}$

Change in magnetization, $\Delta M = M_1 - M_2 = (1591.152 - 1114.24) \times 10^3 = 476.9 \text{ kA/m}$

EXAMPLE 3.2

Find the magnetic field intensity within a magnetic material where

(i) M = 300 A/m and $\mu = 1.5 \times 10-5$ H/m

(ii) **B** = 150
$$\mu$$
T and χ_m = 15

Solution

(i) Given M = 300 A/m and $\mu = 1.5 \times 10^{-5}$ H/m.

Since $\mu = \mu_0 \mu_r$, the relative permeability μ_r is

$$\mu_r = \frac{\mu}{\mu_0} = \frac{1.5 \times 10^{-5}}{4\pi \times 10^{-7}} = 11.94$$

The magnetic field intensity and the magnetization are related to each other as,

$$M = \chi_m H = (\mu_r - 1)H$$

Hence, the magnetic field intensity is

$$H = \frac{M}{\mu_r - 1} = \frac{300}{11.94 - 1} = 27.42 \text{ A/m}$$

(ii) Given $B = 150 \ \mu\text{T}$ and $\chi_m = 15$. The magnetic flux density is $B = \mu H = \mu_0 \mu_r H$.

Hence, the magnetic field intensity is

$$H = \frac{B}{\mu_0 \mu_r} = \frac{B}{\mu_0 (1 + \chi_m)}$$
 (since $\mu_r = 1 + \chi_m$)
= $\frac{150 \times 10^{-6}}{4\pi \times 10^{-7} \times (1 + 15)} = 7.46$ A/m

EXAMPLE 3.3

A magnetic material has a magnetization of 2300 $\text{A} \cdot \text{m}^{-1}$ and produces a flux density of 0.00314 Wb $\cdot \text{m}^{-2}$. Calculate the magnetizing force and the relative permeability of the material.

Solution Given $M = 2300 \text{ A} \cdot \text{m}^{-1}$ and $0.00314 \text{ Wb} \cdot \text{m}^{-2}$.

We know that $B = \mu_0(M + H)$

(i) Magnetic force,
$$H = \left[\frac{B}{\mu_0} - M\right] = \frac{0.00314}{4\pi \times 10^{-7}} - 2300 = 200 \text{ A} \cdot \text{m}^{-1}$$

(ii) Susceptibility,
$$\chi_m = \frac{M}{H} = \frac{2300}{200} = 11.5$$

Hence, relative permeability $\mu_r = 1 + \chi_m = 12.5$

EXAMPLE 3.4

The saturation magnetic induction of Nickel is 0.65 $Wb \cdot m^{-2}$. If the density of Nickel is 8906 kgm⁻³ and its atomic weight is 58.7, calculate the magnetic moment of the Nickel atom in Bohr magneton. (AU Dec 2001)

Solution Given $B_a = 0.65 \text{ Wb} \cdot \text{m}^{-2}$, $\rho = 8906 \text{ kgm}^{-3}$ and atomic weight of Ni = 58.7.

No. of Nickel atoms per m³,
$$N = \frac{\rho \times \text{Avogardro number}}{\text{Atomic weight}} = \frac{8906 \times 6.025 \times 10^{26}}{58.7} = 9.14 \times 10^{28} \text{ m}^{-3}$$

Saturation magnetization, $B = N\mu_0 \mu_m$

Hence, magnetic moment, $\mu_m = \frac{B}{N\mu_0} = \frac{0.65}{9.14 \times 10^{28} \times 4\pi \times 10^{-7}} = 5.66 \times 10^{-24} \,\text{A} \cdot \text{m}^2$

We know that 1 Bohr magneton, $\mu_B = 9.27 \times 10^{-24} \text{ A} \cdot \text{m}^{-2}$

Magnetic moment in terms of Bohr magneton is

$$\mu'_m = \frac{\mu_m}{\mu_B} = \frac{5.66 \times 10^{-24}}{9.27 \times 10^{-24}} = 0.61$$
 Bohr magneton

EXAMPLE 3.5

A magnetic field of 2000 $A \cdot m^{-1}$ is applied to a material which has a susceptibility of 1000. Calculate the (i) magnetization and (ii) flux density.

Solution Given $H = 2000 \text{ A} \cdot \text{m}^{-1}$ and $\chi_m = 1000$.

- (i) Magnetization, $M = \chi_m H = 2000 \times 1000 = 2 \times 10^6 \text{ A} \cdot \text{m}^{-1}$
- (ii) Flux density, $B = \mu_0 (M + H) = 4\pi \times 10^{-7} \times (2 \times 10^6 + 2000) = 2.514 \text{ Wb} \cdot \text{m}^{-2}$

EXAMPLE 3.6

The magnetic field strength of Silicon is 1500 A \cdot m⁻¹. If the magnetic susceptibility is -0.3×10^{-5} , calculate the magnetization and flux density in Silicon.

Solution Given $H = 1500 \text{ A} \cdot \text{m}^{-1}$ and $\chi_m = -0.3 \times 10^{-5}$.

- (i) Magnetization, $M = \chi_m H = 1500 \times -0.3 \times 10^{-5} = -4.5 \times 10^{-3} \text{ A} \cdot \text{m}^{-1}$
- (ii) Flux density, $B = \mu_0 (M + H) = 4\pi \times 10^{-7} \times (-4.5 \times 10^{-3} + 1500) = 1.88 \times 10^{-3} \text{ Wb} \cdot \text{m}^{-2}$

EXAMPLE 3.7

A magnetic field of 1800 ampere/meter produces a magnetic flux of 3×10^{-5} Wb in an iron bar of cross sectional area 0.2 cm². Calculate the permeability.

Solution Given $H = 1800 \text{ A} \cdot \text{m}^{-1}$, $\phi = 3 \times 10^{-5} \text{ Wb}$ and $A = 0.2 \times 10^{-4} \text{ m}^2$.

Flux density,

$$B = \frac{\phi}{A} = \frac{3 \times 10^{-5}}{0.2 \times 10^{-4}} = 1.5 \,\mathrm{Wb} \cdot \mathrm{m}^{-2}$$

Permeability,

$$\mu = \frac{B}{H} = \frac{1.5}{1800} = 8.333 \times 10^{-4} \text{ H/m}$$

3.6 TYPES OF MAGNETIC MATERIALS

Magnetic materials are broadly classified into two types. They are (a) phenomenological classification based on the sign and magnitude of magnetic susceptibility and (b) Microscopic classification based on effects of interaction of the material with external field.

(a) Phenomenological Classification of Magnetic Materials

The phenomenological classification of magnetic materials is divided into three types as shown in Fig. 3.5. This approach ignores the nature of microscopic carriers of magnetism and does not consider their interaction with the external field. Only a different approach at microscopic level can explain the magnetic states like ferromagnetic, anti-ferromagnetic and ferrimagnetic. It was first proposed by Ampere that the magnetic properties of a material arise due to large number of tiny current loops within the material. These tiny microscopic current loops are associated with the motion of electrons within the atoms and each current loop has a magnetic moment associated with it.



Fig. 3.5 Phenomenological classification of magnetic materials

In addition to the orbital motion of electrons around the nucleus, electrons also spin or rotate about their own axis. Thus, the internal magnetic field in a material is produced by electron orbiting around the nucleus and by the spin of the electrons. This is how internal magnetism is produced in the material.

3.10 Physics for Electronics Engineering and Information Science

(b) Microscopic Classification of magnetic materials

The microscopic classification of the magnetic materials is shown in Fig. 3.6.



Fig. 3.6 Microscopic classification of the magnetic materials

The various types of magnetic materials are described below.

3.6.1 Diamagnetism

Diamagnetism is exhibited by materials whose atoms contain even number of electrons. The electrons of such atoms are paired. The electrons in each pair have orbital motions as well as spin motions in opposite sense. Thus the resultant magnetic dipole moment of the atom is zero. The electron pair and hence the atoms acquire an effective magnetic dipole moment which is opposite to the applied field. Hence, for diamagnetic materials, the magnetic moment is opposite to magnetization and the susceptibility of a diamagnetic substance is negative and very small. It is a weak effect and in solids, it is often masked by other kinds of magnetism.

Properties of Diamagnetic Materials

- 1. Permanent dipoles are absent and hence, permanent magnetic moment of each atom is zero as shown in Fig. 3.7.
- 2. In the presence of an external field, they become weakly magnetized but in a direction opposite to the magnetizing field as shown in Fig. 3.8.





Fig. 3.9 Behaviour in presence of field for (a) normal and (b) diamagnetic material

- 4. The relative permeability μ_r is less than 1.
- 5. The susceptibility is independent of temperature and negative, which is of the order of 10^{-5} for most diamagnetic materials such as organic materials and light elements as shown in Fig. 3.10.



Fig. 3.10 Characteristics of diamagnetic material (a) M vs H and (b) χ_m vs T

- 6. In materials like Alkali earths and Bismuth, χ_m is intermediate negative and varies with field and temperature below 20 K.
- 7. Perfect diamagnetism occurs ($\chi_m = -1$) in some materials called superconductors below critical temperatures (Meissner effect) and hence, in superconductors, the magnetic induction \overline{B} will be zero.
- 8. Some other examples of diamagnetic materials are, Antimony, Copper, Zinc, Silicon, Germanium, Graphite, Sulphur, Diamond, Sodium Chloride, Hydrogen, Helium, Gold, Silver, Mercury and Lead.

3.6.2 Paramagnetic Materials

In paramagnetic materials, the magnetic fields associated with orbiting and spinning electrons do not cancel each other due to the presence of unpaired electrons. Each electron in an orbit has an orbital and a spin magnetic moment. When the shells are unfilled, there is a net magnetic moment. In the absence of external magnetic field, the net moments of the atoms are arranged in random directions because of thermal fluctuations. Hence, there is no magnetization. When an external magnetic field is applied, there is a tendency for the dipoles to align with the field giving rise to an induced positive dipole moment proportional to the field. The magnetic moment and the magnetic field are in the same direction and hence, the susceptibility is positive.

When a paramagnetic material is heated, the thermal agitation of its atoms increases and hence, the alignment of the dipoles becomes disturbed. Hence, the magnetization of paramagnetic substances decreases as the temperature increases. In this case, susceptibility is inversely proportional to temperature. For most paramagnetic materials like air, oxygen, platinum, tungsten, potassium, erbium chloride, neodymium oxide and yttrium oxide, χ_m is of the order of 10^{-5} to 10^{-3} .

Properties of paramagnetic materials

1. Paramagnetic materials possess permanent magnetic dipoles which are randomly oriented in the absence of an external applied field. Hence, the net magnetization in any given direction is zero as shown in Fig. 3.11.



Fig. 3.11 Random arrangement of dipoles in the absence of field

2. In the presence of field, the-dipoles align with the field giving rise to an induced positive dipole moment as shown in Fig. 3.12.



Fig. 3.12 Alignment of dipoles in the presence of field

3. When placed inside a magnetic field, it attracts the magnetic lines of force as shown in Fig. 3.13.



Fig. 3.13 Behaviour in presence of field for (a) normal and (b) paramagnetic material

- 4. In alkali and transition metals, susceptibility is small, positive and it is independent of temperature.
- 5. In rare earths, the paramagnetic susceptibility is large and positive and obeys the Curie law, $\chi = \frac{C}{T}$ or the Curie Weiss law, $\chi = \frac{C}{T \theta}$ where *C* is Curie constant and θ is Curie temperature as shown in Fig. 3.14.



Fig. 3.14 Characteristics of paramagnetic material (a) M vs H and (b) χ_m vs T

6. Paramagnetic susceptibility is independent of the applied magnetic field.

3.6.3 Ferromagnetic Materials

Ferromagnetic materials like Iron, Cobalt, Nickel and their alloys are very strongly magnetic. A ferromagnetic material has a spontaneous large magnetic moment i.e., a magnetic moment which is non-zero even in the absence of an external field. This is because they have a net intrinsic magnetic dipole moment due to the spin of the electrons.

Properties of Ferromagnetic Materials

1. The spins align in parallel in the same direction due to a special form of interaction called exchange coupling between adjacent atoms as shown in Fig. 3.15.



Fig. 3.15 Parallel spin alignment in a ferromagnetic material

2. When the ferromagnetic material is placed inside a magnetic field, it attracts the magnetic lines of forces very strongly as shown in Fig. 3.16.



Fig. 3.16 Behaviour in presence of field for (a) normal and (b) ferromagnetic material

3. Each ferromagnetic material has a characteristic temperature called the *ferromagnetic Curie* temperature θ_f . Above θ_f , its properties are quite different from those below. Materials like iron, cobalt, nickel and gadolinium have very large positive ferromagnetic susceptibility which greatly depends on temperature, in accordance with $\chi_m = \frac{C}{T - \theta}$. Here, paramagnetic behaviour predominates for $T > \theta_f$ and ferromagnetic behaviour predominates for $T < \theta_f$ as shown in Fig. 3.17.



Fig. 3.17 Characteristics of ferromagnetic material (a) χ_m vsT and (b) reciprocal of χ_m as a function of T

4. Exhibits hysteresis below θ_{f} . When a ferromagnetic material is taken through a cycle of magnetization, it is seen that the magnetic induction lags behind the applied field as shown in Fig. 3.18. This lagging of \overline{B} behind \overline{H} is called hysteresis.



Fig. 3.18 Schematic representation of the hysteresis loop for a ferromagnetic material

5. These materials consist of a number of small regions which are spontaneously magnetized called domains.

3.6.4 Antiferromagnetism

In ferromagnetism, the tendency for parallel alignment of the electron spins is due to quantum mechanical exchange forces. In certain materials when the distance between the interacting atoms is small, the exchange forces produce a tendency for anti parallel alignment of electron spins of neighboring atoms. This kind of interaction is encountered in antiferromagnetic and ferrimagnetic materials.

The most characteristic feature of an antiferromagnetic material is the occurrence of a sharp maximum in the susceptibility versus temperature curve as shown in Fig. 3.20 (a) for MnF_2 . The temperature at which the maximum occurs is called the Neel temperature, T_N . Above Neel temperature, the susceptibility varies as per the equation.

$$\chi_m = \frac{C}{T + \theta} \tag{3.8}$$

where C is the Curie constant and θ is the paramagnetic Curie temperature

Properties of Antiferromagnetic materials

1. Electron spins of neighboring atoms are aligned anti-parallel as shown in Fig. 3.19.



Fig. 3.19 Spin alignment (a) in general and (b) in a unit cell in an antiferromagnetic material

- 2. Salts of some transition elements exhibit anti-ferromagnetic susceptibility which is small and positive and greatly depends on temperature i.e., $\chi_m = \frac{C}{T+\theta}$ when $T > T_N$ and $\chi_m \propto T$ when $T < T_N$.
- 3. Initially, susceptibility increases slightly with temperature and beyond Neel temperature, the susceptibility decreases with the temperature as shown in Fig. 3.20.



Fig. 3.20 Characteristics of antiferromagnetic material (a) χ_m of MnF_2 vsT (polycrystalline material as a function of temperature) and (b) reciprocal of χ_m as a function of T

3.6.5 Ferrimagnetism

This is a special case of antiferromagnetism. The net magnetization of magnetic sub-lattices is not zero since anti parallel moments are of different magnitudes. Hence, a ferrimagnetic material possesses a net magnetic moment. This moment disappears above the Curie temperature θ analogous to the Neel temperature. Above θ , thermal energy randomizes the individual magnetic moments and the material becomes paramagnetic.

Properties of Ferrimagnetic Materials

- 1. Ferrimagnetic materials possess net magnetic moment as the spin alignment is anti parallel with different magnitudes as shown in Fig. 3.21.
- 2. Above Curie temperature, it becomes paramagnetic while below it behaves as a ferromagnetic material. Ferrites have a susceptibility which is positive and very large. It depends on temperature i.e.,

$$\chi_m = \frac{C}{T \pm \theta}$$
 for $T > T_N$ and is paramagnetic and for $T < T_N$, it is

Fig. 3.21 Spin alignment in a ferrimagnetic material

ferromagnetic as shown in Fig. 3.22.



Fig. 3.22 Graph of χ_m vs T for a ferrimagnetic material

The classifications of all types of magnetic materials are shown in Fig. 3.23(a). The variation of reciprocal of susceptibility in magnetic materials with temperature is shown in Fig. 3.23(b).



Fig. 3.23(a) Variation of flux density with applied field H

(b) Reciprocal of χ_m vs T

The comparison	of magnetic	materials is	tabulated in	n Table 3.1.

Definition and Properties	Diamagnetic material	Paramagnetic material	Ferromagnetic material
Definition	It is a material in which there is no permanent dipole moment or magnetic moment in each atom. The induced magnetic moment produced in these materials during the application of the external magnetic field decreases the magnetic induction present in the specimen.	It is a material in which there is permanent dipole moment or magnetic moment in each atom. The induced magnetic moment produced in these materials during the application of the external magnetic field increases the magnetic induction present in the specimen.	It is a material in which there is enormous permanent dipole moment or magnetic moment in each atom. The induced magnetic moment produced during the application of the external magnetic field is very large and it increases the magnetic induction present in the specimen.
Origin	Arises from the Larmor precession of the electronic orbits in the presence of applied magnetic field.	Arises from the magnetic moments orientation along the external magnetic field direction and magnetic moments orientation is largely determined by temperature and applied magnetic field.	Arises from the spontaneous magnetization due to local molecular magnetic field which arises from exchange interaction between unpaired electrons of adjacent atoms in the crystal lattice.

Table 3.1	Comparison of	different types of	of magnetic	materials
-----------	---------------	--------------------	-------------	-----------

(Contd.)

Definition and Properties	Diamagnetic material	Paramagnetic material	Ferromagnetic material
	No spin or magnetic moment		$ \begin{array}{c} \uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \uparrow $
Spin alignment or magnetic moment alignment		All spins or magnetic moments are randomly oriented.	All spins (or) magnetic moments are orderly oriented. Antiferro magnetic materials (A.M) and ferrimagnetic materials (F'.M) are the special cases of ferromagnetic materials (F.M) on the basis of spin alignment.
Behaviour of material in the presence of			
presence of magnetic field.	In a normal material, $B_{in} = B_{out}$. But in a diamagnetic material, the magnetic lines of force are repelled away from the centre of the material and hence, $B_{in} < B_{out}$	The magnetic lines of force are attracted towards the centre of the material and hence, $B_{in} > B_{out}$	The magnetic lines of force are highly attracted towards the centre of the material and hence, $B_{in} >> B_{out}$
	It is always negative and independent of the temperature and strength of applied magnetic field also.	It is always positive and small and inversely proportional to absolute temperature of the material.	It is always positive and large and dependent on temperature in a complex manner.
Susceptibility and its temperature dependence	χ_m $\chi_m = \text{constant}$	$\chi_m \wedge \frac{1}{T} \rightarrow T$	Ferro region θ_f T
Magnetic phase transition	At very low temperatures $(\approx 1 \text{K})$, the superconductors are perfect diamagnetic. When temperature is increased beyond its critical temperature, diamagnetism suddenly disappears and then it behaves like a normal material.	When the temperature of material is less than its curie temperature, it is converted into diamagnetic substance. But paramagnetic curie temperature is very low (100K)	When the temperature of material is greater than its curie temperature, it is converted into paramagnetic. The ferromagnetic curie temperature is very high.

3.18 *Physics for Electronics Engineering and Information Science*

(Contd.)

Definition and Properties	Diamagnetic material	Paramagnetic material	Ferromagnetic material
Examples	Hydrogen, Bismuth and superconducting materials like Niobium and its components.	Aluminium, Platinum, crown glass and Iron (or) Nickel salt solutions.	Ferromagnetic materials: Iron, Nickel and Cobalt. Antiferromagnetic materials: Ferrous oxide, Zinc ferrite and Manganese oxide. Ferrimagnetic materials: Nickel ferrite, Manganese ferrite and Ferrous ferrite.

EXAMPLE 3.8

A paramagnetic material has BCC structure with the cube edge of 2.5 Å. If the saturation value of magnetization is $1.8 \times 10^6 \text{ A} \cdot \text{m}^{-1}$, calculate the average magnetization contributed per atom in Bohr magneton.

Solution Given $a = 2.5 \times 10^{-10}$ m, $M = 1.8 \times 10^{6}$ A·m⁻¹ and n = 2 (for BCC structure).

No. of atoms per unit volume = $\frac{\text{No. of atoms present in an unit cell}}{\text{Volume of the unit cell}}$

i.e.,

$$N = \frac{n}{a^3} = \frac{2}{(2.5 \times 10^{-10})^3} = 1.28 \times 10^{29} \text{ atoms/m}^3$$

Average magnetization produced per atom is

$$M' = \frac{M}{N} = \frac{1.8 \times 10^{\circ}}{1.28 \times 10^{29}} = 1.4062 \times 10^{-23} \,\mathrm{A \cdot m^{-2}}$$

Bohr magneton, $\mu_B = \frac{eh}{4\pi m_e} = \frac{1.6 \times 10^{-19} \times 6.625 \times 10^{-34}}{4 \times 3.14 \times 9.1 \times 10^{-31}} = 9.27 \times 10^{-24} \,\mathrm{A \cdot m^{-2}}$

Hence, average magnetization produced per atom in Bohr magneton is

$$\frac{M'}{\mu_B} = \frac{1.4062 \times 10^{-23}}{9.27 \times 10^{-24}} = 1.517$$
 Bohr magneton

EXAMPLE 3.9

The susceptibility of paramagnetic FeCl₃ is 3.7×10^{-3} at 27° C. What will be the value of its susceptibility at 200 K and 500 K?

Solution Given $\chi_m = 3.7 \times 10^{-3}$.

At room temperature = 300 K, Curie constant is given by,

$$C = \chi_m T = 3.7 \times 10^{-3} \times 300 = 1.11 \text{ K}$$

(i) At 200 K,
$$\chi_m = \frac{C}{T} = \frac{1.11}{200} = 5.55 \times 10^{-3}$$

(ii) At 500 K,
$$\chi_m = \frac{C}{T} = \frac{1.11}{500} = 2.22 \times 10^{-3}$$

3.7 FERROMAGNETISM – ORIGIN AND EXCHANGE INTERACTION

As discussed earlier, ferromagnetic materials like iron exhibit permanent magnetization even in the absence of an applied field, i.e., they possess a susceptibility which is infinite.

In a magnetized iron crystal, all the atomic magnetic moments are aligned in the same direction, as shown in Fig. 3.24 (a). As the magnetic potential energy of interaction is smaller than the thermal energy, there cannot be alignment of the moments and the resulting large magnetization cannot be due to the magnetic forces between the moments. This large value of magnetization has been explained by Heisenberg as due to quantum exchange interactions between electrons which are electrostatic in nature. A detailed discussion is given below.



Fig. 3.24 (a) Isolated Fe atom having four unpaired spins and (b) Hund's rule for an atom with many electrons

Iron, cobalt, and nickel fall under the category of 3d transition elements which have unpaired electrons and are ferromagnetic at room temperature. The rare earth metals gadolinium and dysprosium are ferromagnetic below room temperature.

The iron atom has the electronic configuration $3d^64s^2$. In an isolated iron atom, only four 3d sub shells are filled with unpaired electrons as shown in Fig. 3.24 (b). The isolated atom has four parallel electron spins and hence a spin magnetic moment of $4\mu_B$.

The origin of Hund's rule, visualized in Fig. 3.24 (b), lies in the fact that when the spins are parallel (same spin quantum number m_s), as a requirement of the Pauli exclusion principle, the electrons must occupy orbitals with different m_l (m_l , the azimuthal quantum number determines the orientation of an orbit). This results in a reduction of Couloumbic repulsion between the electrons. Consequently, the electrostatic interaction energy is minimized. This constitutes an exchange interaction between 2 electrons with parallel spins. Hence, in an atom, this exchange interaction, forces two electrons to take the same m_s but different m_l values, without violating the Pauli exclusion principle. These exchange interactions are also called spin dependent electrostatic forces and are much greater than the magnetic interaction forces. This is the reason an isolated Fe atom has four unpaired spins in the 3d subshell.

Similarly, isolated Cobalt and Nickel atoms have spin magnetic moments of $3\mu_B$ and $2\mu_B$ respectively.

In solids, the electron spins are affected by proximity of other atoms. The number of electrons that actually parallel their spins depends on the strength of the exchange interaction, and the actual magnetic moments are much smaller. These values are $2.22\mu_B$, $1.77\mu_B$ and $0.606\mu_B$ for Iron, Cobalt and Nickel respectively.

In the simplest case, for two atoms *i* and *j*, the exchange energy depends on the inter-atomic separation between two interacting atoms and the relative spins of the two outer electrons. From quantum mechanics, the exchange interaction can be represented in terms of exchange energy E_{ex} as

$$E_{\rm ex} = -2J_{ij}S_iS_j \tag{3.9}$$

where S_i and S_j are the spin angular momenta of the two electrons and J_{ij} is a numerical quantity called the exchange integral that involves integrating the wave functions with the various potential energy interaction terms.

The total exchange interaction energy between spins of unpaired electrons of the system is given by

$$E_{\text{ex}} = -2\sum_{ij} (J_{ij})_e \overrightarrow{S_i} \cdot \overrightarrow{S_j}$$
(3.10)

The value of J_{ii} is

$$J_{ij} = \int \psi_A^*(1) \psi_B^*(2) V_{AB} \psi_A(1) \psi_B(2) dr_1 dr_2$$
(3.11)

which gives the probability that electron 1 in atom A will exchange with electron 2 in atom B ψ_A and ψ_B are the wave functions for atoms A and B respectively and 1 and 2 represent the unpaired electrons in A and B respectively.

The interaction potential V_{AB} between ions A and B and electrons 1 and 2 as given by

$$V_{AB} = \frac{e^2}{4\pi\varepsilon_0} \left[\frac{1}{r_{AB}} + \frac{1}{r_{12}} - \frac{1}{r_{B1}} - \frac{1}{r_{B2}} - \frac{1}{r_{A1}} - \frac{1}{r_{A2}} \right]$$
(3.12)

where r_{AB} is the distance between atoms A and B and r_{12} is the distance between electrons 1 and 2 and so on.

Therefore, it depends on the electrostatic interactions and hence on the inter atomic distance. For majority of solids, J_{ii} is negative. This is the antiferromagnetic state. For Fe, Co, and

Ni, J_{ij} is positive. E_{ex} is then negative if S_i and S_j are parallel. Spins of the 3d electrons of the Fe atoms spontaneously align in the same direction to reduce the exchange energy. This spontaneous magnetization is the phenomenon of ferromagnetism. Figure 3.25 shows how J_{ij} changes with the ratio of inter atomic separation to the radius of the 3d sub shell (r/r_d) . The corresponding values are shown in Table 3.2. For the transition metals like Fe, Co, and Ni, the value of $(r/r_d) > 3$ and J_{ij} is positive. Hence, exchange energy is negative for parallel orientation of spins and the material exhibits ferromagnetism. In all other cases, J_{ij} is negative and does not produce ferromagnetic behaviour.



Fig. 3.25 Exchange integral as a function of r/r_d

Table 3.2 Values of $((r/r_d))$ for different atoms

	Fe	Со	Ni	Cr	Mn	Gd
$((r/r_d))$	3.26	3.64	3.94	2.60	2.94	4.12

It is possible to combine two materials with $(r/r_d) > 3$ and $(r/r_d) < 3$ to form new compounds for which $(r/r_d) > 3$ and thus, make it ferromagnetic. Mn-As, Cu-Mn and Mn-Sb alloys are some examples.

3.8 SATURATION MAGNETIZATION AND CURIE TEMPERATURE

The maximum magnetization in a ferromagnet when all the atomic magnetic moments have been aligned as much as possible is called the saturation magnetization M_{sat} . With increase in temperature, lattice vibrations increase and disrupt the alignments of the spins resulting in random orientation of spins. The randomness increases till a critical temperature called the Curie temperature, denoted by θ_{f} , when the thermal energy of lattice vibrations in the crystal can overcome the potential energy of the exchange interaction and hence, destroy the spin alignments. For example, in a ferromagnetic material, the ferromagnetic behaviour disappears at Curie temperature and above the Curie temperature, the crystal behaves as paramagnetic. The saturation magnetization M_{sat} , decreases from its maximum value $M_{sat}(0)$ at absolute zero of temperature to zero at the Curie temperature. Figure 3.26 (a) shows the dependence of M_{sat} on the temperature for Iron, when M_{sat} has been normalized to $M_{sat}(0)$ and the temperature is the reduced temperature, i.e., T/θ_{f} .

At $T/\theta_f = 1$, $M_{sat} = 0$. When plotted in this way, the ferromagnets like cobalt and nickel follow closely the observed behaviour for iron as shown in Fig. 3.24(b). It is noted that for iron, $\theta_f = 1043$ K at room temperature, $T/\theta_f = 0.29$ and M_{sat} is very close to its value at $M_{sat}(0)$.



Fig. 3.26 Magnetization vs temperature (a) Normalized saturated magnetization vs reduced temperature T/ θ_f and (b) Variation of relative magnetization for Fe, Co and Ni as a function of T/ θ_f

Since at the Curie temperature, the thermal energy, of the order of $k\theta_f$, is sufficient to overcome the energy of the exchange interaction E_{ex} that aligns the spins, θ_f is taken as the order of magnitude of E_{ex} . For iron, E_{ex} is ~0.09 eV and for cobalt, it is ~0.1 eV. Table 3.3 summarizes some of the important properties of the ferromagnets Fe, Co, Ni, and Gd (rare earth metal).

Table 3.3	Properties of the	ferromagnets Fe	e, Co	, Ni and Gd
				/

Properties	Fe	Со	Ni	Gd
Crystal structure	BCC	НСР	FCC	НСР
Bohr magnetons per atom (Am^2)	2.22	1.72	0.60	7.1
$M_{sat}(0)(MAm^{-1})$	1.75	1.45	0.50	2.0
$B_{sat} = \mu_0 M_{sat}(T)$	2.2	1.82	0.64	2.5
θ_{f} (Ferromagnetic Curie temperature)	770°C 1043K	1127°C 1400K	358°C 631K	16°C 289K

From the parallel spin alignment due to exchange interactions and saturation magnetisation, it would seem that a single crystal of iron should possess a net permanent or spontaneous magnetisation in the absence of an applied field even at room temperature (since $\theta_f = 1043$ K for iron).

But this is not the actual case. If a magnetised piece of iron is heated to a temperature above its Curie temperature and allowed to cool in the absence of a magnetic field, it will possess no net magnetisation. Further, in the presence of field, it exhibits Hysteresis.

According to Weiss theory, a ferromagnetic solid is divided into a large number of small regions called domains, each of which is spontaneously magnetised to saturation. Within a domain, all spins are aligned parallel so as to produce a magnetic moment in one direction only. Since different domains are randomly oriented, the magnetic moments cancel each other resulting in the net magnetisation to become zero.

3.9 DOMAIN THEORY OF FERROMAGNETISM

In order to explain the absence of spontaneous magnetisation in a virgin ferromagnetic material, knowledge about domain formation in terms of energy is important. It is discussed in detail in this section.

3.9.1 Formation of Domains

Every material tries to attain a stable form by minimizing its energy. The domain structure of a material is a one that minimizes the Gibbs free energy of the material and is given by the free energy equation proposed by Landau and Lifshitz in 1935. This free energy E_g (which is to be minimized) is given by

$$E_g = E_{\text{ex}} + E_{\text{ms}} + E_{\text{ani}} + E_{\text{mst}} + E_Z$$

where E_{ex} is the exchange energy, E_{ms} is the magnetostatic energy, E_{ani} is the anisotropic energy, E_{mst} is the magnetostrictive energy and E_Z is the Zeeman energy which comes into play only in the presence of an external field and hence, will not be discussed here. Also some times $E_{ex} + E_{ani}$ is referred to as E_{bw} (the Bloch wall energy). Each type of energy is described below.

(a) Exchange Energy

Some ferromagnetic materials like Fe, Co and Ni, have very high value of magnetic moments due to the presence of unpaired electrons as shown in Fig. 3.27.

This arises due to strong quantum mechanical exchange interactions between dipoles and is called *exchange energy*. This will be less for parallel spin alignment as shown in Fig. 3.28 (a) when compared to antiparallel spin alignment. Hence, all the dipoles will first align in the same direction creating north and south poles due to its self-magnetic moment as shown in Fig. 3.28 (b).







Fig. 3.28 Exchange energy (a) showing parallel spin alignment and (b) net magnetic moment due to parallel spins

(b) Magnetostatic Energy

But due to the above, magnetic lines of force will be created outside the material as shown in Fig. 3.29. This self-energy created by magnetization in some parts of the sample on the other parts is called *magnetostatic energy* and depends on the volume occupied by the magnetic field extending outside the domain and the sample.

(c) Anisotropic energy

By dividing a single domain into N domains, the magnetostatic energy can be reduced to $1/N^{\text{th}}$ of the magnetic energy of that of a single domain and is shown in Figs. 3.30 (a) and (b). This is done by the spins changing their direction thereby, creating a Bloch wall.

For this, the spins have to shift from easy to hard directions of magnetization and require energy to be spent. From Fig. 3.31, it can be seen that a stronger field is required to saturate BCC iron (Fe) in the hard direction. As the dipoles will naturally be aligned in easy directions, some work has to be done to turn them into the hard direction and attain the configuration as shown in Fig. 3.30(b). This difference between the energy to produce saturation in the easy and hard directions is called *crystal anisotropy energy*.



Fig. 3.29 External magnetic lines of force



3.30 Reduction of magnetostatic energy by division into (a) two domains and (b) four domains



Fig. 3.31 Easy and hard directions of magnetization in BCC Fe

(d) Bloch Wall Energy

The process of changing the directions of spins from one to another can be either gradual or abrupt and this creates a Bloch wall as shown in Fig. 3.32. If the change is gradual, exchange energy is lower but anisotropy energy is larger. If the change is abrupt, exchange energy is larger but anisotropy energy is lesser. The Bloch wall is a compromise between the two. The wider the wall is, the greater will be the number of dipoles forced to line in directions different from those of easy magnetization, and the magneto crystalline anisotropy energy will be increased. Hence, the wall width will reach an equilibrium value when,

$$E_{\rm ex} + E_{\rm ani} = E_{\rm bw}$$
 (Minimum)

From the above equation, it is known that the domain wall formation requires extra energy called "Domain wall energy" which is proportional to the area of the domain wall (square of domain size). But the field energy is proportional to volume of domains (cube of domain size). As domains become smaller, the net energy saved by splitting decreases because the net energy reduced is the difference between the magnetic field energy saved and the additional energy required to create the domain wall. The domains will go on splitting into smaller and smaller regions till the energy required to create an additional domain wall just equals the field energy saved. Only such domain sizes will be stable. This is of the order of 10^{-4} to 10^{-6} m in most materials.



Fig. 3.32 (a) Bloch wall formation and (b) graph showing minimum Bloch energy

(e) Magnetostrictive Energy

Finally, flux closure domains will form with magnetization at right angles to other domains as shown in Fig. 3.33. This reduces the magnetostatic energy to nearly zero but increases the anisotropy energy and also causes small mechanical stresses in the material. This is termed as *magnetoelastic anisotropy* or simply *magnetostrictive energy*. Hence, closure domains formation require additional energy and will be formed only when the magnetostatic energy saved is more than the sum of the exchange energy, magnetocrystalline anisotropy energy and magnetostrictive energy.

i.e.,
$$E_{\rm ms} \text{ saved} > E_{\rm ex} + E_{\rm ani} + E_{\rm mst}$$



Fig. 3.33 Closure domain structure where magnetostatic energy is zero

Finally, the size of the domain will be stable when the free energy $E_g = E_{ex} + E_{ms} + E_{ani} + E_{mst}$ attains a minimum. To achieve the above, E_{ex} , $E_{ani} \& E_{mst}$ increase proportionately to reduce E_{ms} .

3.9.2 Grain Structure

The magnetic domains for a perfect lattice like a single crystal of iron were discussed in the previous section. But most materials are polycrystalline. Grains are not same as domains. Each grain is a single crystal and is big enough to contain several domains. Each crystal has "easy" axes of magnetization and is divided into domains with axis of magnetization parallel to these directions. Due to the random orientation of different domains, the net magnetic moment in any direction is almost zero. On application of an external field, the spins in each domain align in the direction of the field. The orientation of domains in the presence and absence of field are shown in Figs. 3.34 (a) and (b).



Fig. 3.34 Domain alignment in the (a) absence of magnetic field and (b) presence of magnetic field

Experimental Evidence of Domains (Bitter Powder Patterns)

The domain structure was observed in colloidal iron oxide to delineate domain boundaries. The photo micrographs are shown in the Fig. 3.35. They are obtained using the bitter colloidal iron oxide technique. They illustrate the growth of favorable domain in iron single crystals when the field is increased. The direct evidence for the existence of domain structure is obtained from these bitter powder patterns. In this technique, a drop of colloidal suspension of finely divided ferromagnetic powder is allowed to spread over a prepared surface of ferromagnetic material under investigation. As observed under microscope, the colloidal particles, collect along the domain boundaries since strong local magnetic fields exist near these domain boundaries. When the sample is kept in the magnetic field, the domain boundaries are shifted depending upon the direction of applied magnetic field. When the intensity of the magnetic field increases, the area of the favorable domains increases by the shifting of domain boundaries as shown in Fig. 3.35.



Fig. 3.35 Bitter powder patterns

3.9.3 Hysteresis

The hysteresis of ferromagnetic materials refers to the lag of magnetization behind the magnetizing field. When the temperature of the ferromagnetic substance is less than ferromagnetic Curie temperature, the substance exhibits hysteresis. The domain concept is well suited to explain the phenomenon of hysteresis. The increase in the value of the resultant magnetic moment of the specimen under the action of the applied field can be attributed to (i) the motion of domain walls and (ii) rotation of domains as shown in Fig. 3.36.



Fig. 3.36 Domain growth (a) in absence of field (b) weak field (c) high field

When a weak magnetic field is applied, the domains that are aligned parallel to the field and in the easy direction of magnetization grow in size at the expense of less favorably oriented ones.



Fig. 3.37 Magnetization curve of a virgin material (a) indicating domain movements in half cycle and (b) full cycle with schematic diagram of domains

This makes the domain wall or the Bloch wall to move. When the field is removed, the domains go back to their initial state. This is shown as part OA in Fig. 3.37(a). If the field is increased further, the domain wall also moves further and the movement becomes irreversible and is shown as path AB in Fig. 3.37(a). On reaching point B, all domains are magnetized along their easy directions. Further increase of field to higher values rotates the domains so as to align along the field direction. This will be away from the easy direction and hence, anisotropy energy will be stored. Once the domain rotation is complete, the specimen is saturated as denoted by C in the Fig. 3.37 (a).

On removal of the field, the specimen tends to attain the original configuration by the movement of Bloch walls. But this movement is hampered by impurities, lattice imperfections etc., and so more energy must be supplied to overcome the opposing forces. This means that a coercive field is required to reduce the

magnetization of the specimen to zero. The amount of energy spent for this is a loss. Hysteresis loss is the loss of energy in taking a ferromagnetic body through a complete cycle of magnetization and this loss is

represented by the area enclosed by the hysteresis loop. Figure 3.37(b) shows the full cycle with schematic diagram of domains. A typical hysteresis loop is shown in Fig. 3.38. An explanation of the same is given below.

Consider a ferromagnetic material whose initial magnetization M = 0. As the magnetizing field \overline{H} is increased, magnetization increases. It reaches a maximum value at point P, which is called saturation magnetization M_s . When the magnetic field is further increased, the magnetization does not increase and remains as M_s . If the field is decreased to zero, the magnetization attains a value M_r called residual magnetization at point Q.



Fig. 3.38 Hysteresis loop of a ferromagnetic material

When the magnetic field is reversed, i.e., reducing it from zero to negative values, the magnetization of material becomes zero at point R, and is referred as the coercivity, H_C of the specimen. When the field is further increased in the reverse direction, the magnetization of the material reaches its peak value at point S.

If the polarity is reversed, the magnetic field is first slowly increased to zero. Then, if the polarity is made

positive, the magnetization slowly decreases first to the residual value, then to zero and finally reaches the original saturation value. This completes one cycle of magnetization. The amount of energy spent in taking the material through one cycle of magnetization is given by the area of hysteresis loop. This is dissipated in the form of heat and is termed as *hysteresis loss*.

3.9.4 Energy Product of a Magnet

The product of retentivity B_r and coercivity H_C is called as energy product and represents the maximum energy that is stored in the material. Figure 3.39 shows a demagnetization curve in the second quadrant of a *B*-*H* graph along with the corresponding *B*-*BH* curve in the first quadrant. For permanent magnets, the energy product should be very high.



3.10 HARD AND SOFT MAGNETIC MATERIALS- EXAMPLES AND USES

The magnetic materials are classified into two types namely hard and soft magnetic materials based on the area of the hysteresis loop as given in Fig. 3.40.



Fig. 3.40 Hysteresis loop of hard and soft magnetic materials

3.10.1 Hard Magnetic Materials

The materials which cannot be easily magnetised or demagnetised are called hard magnetic materials.

Properties of hard magnetic materials

- 1. The area of the Hysteresis loop is large and hence, hysteresis loss is high.
- 2. The coercivity and retentivity are large.
- 3. Permeability and susceptibility are low.
- 4. Since they retain magnetism, they can be used to make permanent magnets.
- 5. Eddy current loss is high.

Ex. Alnico alloy, steel, hard ferrites or ceramic magnets etc.,

The hard magnetic materials have a number of applications which include automotive, telecommunication, data processing, consumer electronics, aerospace, surgical scanners, electronics and instrumentation. The hard magnetic materials are used for making permanent magnets which are used in motors, dynamos, voltmeters, loudspeakers, ammeters and microphones. The important hard magnetic materials, their properties and applications are given in Table 3.4.

Table 3.4	Properties and	applications of	of some hard	magnetic materials
-----------	----------------	-----------------	--------------	--------------------

Name of the alloy	Retentivity (B_r) weber m^{-2}	$Coercivity (H_C) kAm^{-1}$	Other properties	Applications
High carbon steel	0.9	3.98	High strength and hardness	Permanent magnets
Cobalt steel	0.95	18.31	High coercivity and magnetic strength	Permanent magnets
Al-Ni-Co alloys	0.8-1.2	60-120	 High magnetic and mechanical strength Produced by powder metallurgy Alnico V is more powerful than Alnico II due to higher percentage of cobalt 	Powerful permanent small magnets

(Contd.)

Name of the alloy	Retentivity (B_r) weber m^{-2}	Coercivity (H_C) kAm^{-1}	Other properties	Applications
Cunifes	0.53	36	 Carbon free machinable alloy Magnetic properties are improved by hardening 	Large permanent magnets
Cunico	1	40	Carbon free machinable alloy	Large permanent magnets
Barium ferrite	0.2	140	High coercivity and stable magnetic properties	Powerful magnets

3.10.2 Soft Magnetic Materials

Materials which can be easily magnetised or demagnetised are called soft magnetic materials.

Properties of soft magnetic materials

- 1. The area of the Hysteresis loop is small and hence, hysteresis loss is small.
- 2. The coercivity and retentivity are low.
- 3. Permeability and susceptibility are high.
- 4. Since they can be easily magnetized or demagnetized, they are used to make electromagnets.
- 5. Eddy current loss is low due to high resistivity.

Ex: Iron silicon alloys, Ferrous nickel alloy, pure iron, Mu-metals etc.,

The soft magnetic materials are used in AC (Alternating Current) and DC (Direct Current) applications. The soft magnetic materials are used in magnetic data storage devices for reading, writing and storage along with electric door bells, relays, transformer cores. Table 3.5 lists some of the important soft magnetic materials, their properties and applications.

Name of the alloy	Relative permeability (μ_r)	Coercivity (H _C) kAm ⁻¹	Other properties	Applications
Fe-Si alloy	500–7000	1–100	Resistivity is 6 times that of pure iron	Cores of power transformers
Permalloy	2700–25000	0.4	Higher hardness	Cores of audio and video transformers
Supermalloy	100,000-8000,000	0.4	High hardness	Cores of audio and video transformers where low coercivity is required
Ni-Zn ferrite	500–5000	80	High resistivity and narrow hysteresis loop	Cores of audio and TV transformers
Mn-Zn ferrite	2000–2500	7	High dielectric constant	Microwave isolators and gyrators

Table 3.5	Properties and	applications of	some soft ma	anetic materials
				J

3.10.3 Comparison of Hard and Soft Magnetic Materials

The comparison of hard and soft magnetic materials is given in Table 3.6.

 Table 3.6
 Comparison of hard and soft magnetic materials

Hard magnetic material	Soft magnetic material	
	$ \xrightarrow{H} H $	
Fig. T 3.6 (a) Hysteresis loop of steel	Fig. T 3.6 (b) Hysteresis loop of soft iron	
1. Figure T3.6 (a) shows the nature of hysteresis loop of a hard magnetic material.	Figure T3.6 (b) shows the nature of hysteresis loop of a soft magnetic material.	
2. Hard magnetic materials are magnetic materials which cannot be easily magnetized and demagnetized.	Soft magnetic materials are magnetic materials which can be easily magnetized and demagnetized.	
3. They have large hysteresis loss due to large hysteresis loop area.	They have small hysteresis loss due to small hysteresis loop area.	
4. These materials have small values of permeability and susceptibility.	These materials have large values of permeability and susceptibility.	
5. The eddy current loss is more due to its lower resistivity.	The eddy current loss is less due to its higher resistivity.	
6. In a hard magnetic material, the domain wall movement is difficult owing to crystal imperfections and is irreversible in nature.	In a soft magnetic material, the domain walls move easily and reversibly so that magnetization changes by large amounts for small changes in the magnetizing field.	
7. The coercivity and retentivity are large.	The coercivity and retentivity are small.	
8. In these materials, the irregularities (in the crystal structure) like mechanical strains will be more. Its magnetostatic energy is large.	These materials are free from irregularities (in the crystal structure) like strains or impurities. Its magnetostatic energy is very small.	
9. These are produced by heating the material and then plunging it suddenly into cold oil (quenching process) which sets up internal stresses. Hence, mechanical strains are purposely introduced to make it hard magnetic material.	These are manufactured by heating the pure material to a temperature where sufficient movement of the atoms is possible for them to settle into an ordered lattice, followed by a slow cooling (annealing process) so as not to disturb it.	
10.Examples: Alnico alloy, Cunifes, Cunico and Silmanal.	Examples: Iron silicon alloy, ferrous nickel alloy, ferrite and garnets.	
11. These are used to produce permanent magnets. Permanent magnets are used in magnetic detectors, microphones, flux meters, voltage regulators, damping devices and magnetic separators.	These are used in electro-magnetic machinery and in transformer cores. These are used in switching circuits, microwave isolators; shift registers and matrix storage of computers and to produce electromagnets.	

3.11 MAGNETIC PRINCIPLE IN COMPUTER DATA STORAGE

The soft and hard magnetic materials are used in data storage in computers. Ferromagnetic materials are commonly used for storing data because they can be permanently magnetised. The storage of data generally involves the following steps.

- 1. Writing data on a disc or tape (recording medium) which may be in the form of electrical signal.
- 2. Reading back the data whenever required.

But care should be taken to choose a material such that it resists self-demagnetisation for reliability of data storage. Iron (III) oxide is used in hard disk drives as the magnetic material.

The writing/ reading (W/R) heads should be made of a soft magnetic material like Ni-Fe alloy, i.e., materials that can be easily magnetised and demagnetized to avoid the residual magnetism in the material. They essentially consist of an electromagnet i.e., a coil wound around a soft magnetic material. They are mounted right on top of the disc with a small gap of the order of microns. The data to be written is converted into voltage and passes through the coil. This magnetises the writing head which is a soft magnetic material. As the disc coated with a ferromagnetic material contains of large number of domains, each of which contains magnetic spins aligned parallel, they are affected permanently by the magnetic field induced by the head. As the disc rotates, the signals are thus continuously stored in the form of spins, up or down as shown in Fig. 3.41.



Fig. 3.41 Schematic diagram of magnetic tape recording and readout system

In earlier hard discs, the spins were oriented horizontally and parallel to the disk surface. Later, the perpendicular orientation was adopted so that more data could be stored.

To read the data, the reverse mechanism is adopted. The disc is rotated below the read head. The moving disc induces a magnetic field in the read head which in turn creates a current in the coil around it and is again reconverted back to the original signal.

3.12 MAGNETIC HARD DISK (GMR SENSOR)

It is clear from the above explanations, that a write head magnetizes a region by generating a strong local magnetic field, and a read head detects the magnetisation of the regions.

Initially, hard disks used an electromagnet to magnetise a region and then read the magnetic field data by using Faraday's principle of electromagnetic induction. The next version used the inductive heads which included Metal In Gap (MIG) and thin film heads. As the necessity of data storage increases, read heads

using the principle of Magneto Resistance (MR) are evolved. The principle here is that electrical resistance of the head changes according to the strength of the magnetism from the storage regions called platters in the disc. When this field is switched OFF, the material achieves its original resistance.

The principle of spintronics used in read heads are based on Giant Magneto Resistance (GMR). The read and write elements are separate, but very close to each other on the head portion of an actuator arm. The read element is magneto-resistive and the write element is thin-film inductive. The heads are kept extremely close to the platter with a very small air gap which is in the order of microns. The air moves almost at the speed of the platter. The record and playback head are mounted on a slider. There is only a small air gap between the slider and the surface next to the platter so that they are fast out of contact. This helps in storing terabytes of data in computers.

3.12.1 Principle of GMR

The Giant magnetoresistance (GMR) effect was discovered by Albert Fert and Peter Grunberg for which, they won the Nobel Prize in Physics in 2007.

The charges moving in a magnetic field experience a Lorentz force perpendicular to both the magnetic field and the direction of its motion. Also, a charge moving parallel (or antiparallel) to a magnetic field does not experience any magnetic force.

If a magnetic field is present in a current-carrying wire in a direction other than the direction of current flow, the conduction electrons will experience a magnetic force perpendicular to direction of current. As a result, the electrons will be scattered thereby increasing the resistance. This is called magnetoresistance (MR) effect. Thus, a change in the magnetic field will cause a change in the resistance and hence, a change in the current which can be measured. This effect can be used in sensor and magnetic data readers by a change in magnetic field.

It is known from domain theory that the spins of electrons in a ferromagnetic material can be either "spin-up" or "spin-down", based on the component of magnetic field along a chosen axis. When an electron moves into a magnetized region, it gets scattered depending on its spin. Scattering will be lesser for electrons with spins in the direction of the magnetic field as compared to these with spins opposite the direction of the magnetic field as shown in Fig. 3.42. This principle is used in a GMR sensor.



Fig. 3.42 Schematic diagram of spin direction in GMR

In GMR sensors, a non-magnetic material (spacer) is sandwiched between two layers of a ferromagnetic material alternately as shown in the Fig. 3.43. Such a material is called magnetic super-lattice. It is quite difficult to produce magnetic super lattices, but the effect has been seen in three-layer devices called "spin valves". The name spin valve is given due to the fact that only some directions of electron spins are allowed and the others are scattered, which is similar to a valve.

3.34 Physics for Electronics Engineering and Information Science



Fig. 3.43 Schematic diagram of GMR in absence of external field

In the absence of an external field, alternate layers of the ferromagnetic material will naturally align with opposite magnetization for stable configuration as shown in Fig. 3.43.

- All the electrons coming into this region will be scattered as they will have a spin opposite to the magnetization direction in some region.
- When subjected to an external field, the ferromagnetic spins will align in the direction of the external field. Thus only half of the electrons coming into the region, i.e., those with spin opposite to external field will be scattered to a maximum as shown in Fig. 3.44.



Fig. 3.44 Schematic diagram of GMR in presence of external field

Giant magneto resistance is a quantum mechanical magneto resistance effect observed in multilayers made from alternating ferromagnetic and non-magnetic conductive layers. In the presence of a magnetic field in a magnetic super-lattice, scattering of electrons is reduced by almost half, thereby decreasing the resistance enormously. This effect is based on the dependence of electron scattering and hence the resistance on the spin orientation. The changes in resistance can be observed when a current passes through the GMR element and such voltage fluctuations or signals act as a sensor in these magnetic super-lattices. Magnetic field sensors are used to read data in hard disc drives, biosensors, Micro Electro Mechanical Systems (MEMS) and Magneto-resistive Random Access Memory (MRAM).

3.12.2 Parts of a Magnetic Hard Disk

The parts of a magnetic hard disk are shown in Fig. 3.45. The disk surface is divided into concentric tracks. Thin tracks can hold more data. The data bits are recorded as tiny spots on the tracks. The more bits per inch and greater storage can be attained in a thin spot. Even though the outer tracks are physically longer than the inner ones, most disks hold the same number of bits on each track. The bits are packed as tightly as possible within each track. To read or write at a time, the circular tracks are further divided into sectors. A sector can contain several records or one record can span several sectors.



Fig. 3.45 Photograph showing parts of a magnetic hard disc

A hard disk contains one or more shiny silver platters in which data can be stored magnetically. It contains a read – write

head which moves a tiny magnet over the platters to record or store information. This is driven by an arm mechanism. An electronic circuit links the hard drive to the rest of the computer.

- 1. Actuator This moves the read-write arm. These are mostly voice coils and are simple electromagnets which operate similar to the moving coils that make sound in loudspeaker. They are not affected by temperature variations.
- 2. Read-write arm This swings the read-write head back and forth across the platter.
- 3. Central spindle The platter can rotate at high speed around this.
- 4. Magnetic platter Used for storing data in binary form.
- 5. Plug connections For connecting hard drive to circuit board in a PC.
- 6. Read-write head A minute or tiny magnet at tip of the read-write arm.
- 7. Circuit board This is underneath the disk and controls the flow of data to and from the platter.
- 8. Flexible connector Delivers data from the circuit board to read-write head and platter.
- 9. Small spindle Used to swing the read-write arm across the platter.

The most important part of a hard drive is the platter. It is made using glass or aluminium, over which a thin layer of metal is coated and this can be magnetized or demagnetized. A small hard drive consists of only one platter. Larger drives have a series of platters stacked one over other vertically on a central spindle, with a small gap between them. They can be rotated at a speed of 10,000 rotations per minute (rpm) and this helps the read-write heads to access any part of them.

For each platter, there are two read-write heads, one to read the top surface and the other to read the bottom. They are mounted on an electrically controlled arm that moves from the center of the hard disk to the rim and back again. A layer of fluid or air between the head and the platter surface helps to reduce the wear and tear because of the contact between the head and platter.

Reading and Writing Data

In order to access data from a computer, data has to be stored in an orderly manner on each platter. The read and write process is carried out in a systematic way as follows. A Platter consists of circular paths called

tracks. Data can be arranged in the form of bits in a concentric manner along these paths. The tracks are further divided into smaller portions called sectors as shown in Fig. 3.46. Sectors which have been used and those that are free are stored as maps by part of the hard drive. To store new information, the free sectors are identified and subsequently the read-write head moves across the platter to exactly the right location and writes the data. For data retrieval, the reverse process is adopted.



Fig. 3.46 Platter assembly showing tracks and sectors

An electronic device has to manipulate all the data stored in a hard drive. The interface links the hard disk and the controller. Controllers are either built into the circuit board of the disk drive or into the mother board of the computer.

The advantages of using magnetic hard disks are; a lot of information can be stored in a tiny space, they have longer lifespan, they utilize a limited amount of space, they have high-speed accessibility and are cheap to make.

The disadvantages are the dust or dirt plays a spoilsport if they get into hard discs. A tiny piece of dust makes the read-write head bounce up and down. This makes it crash into the platter and damages its magnetic material. This is called disk crash (or head crash). This may cause a loss of information of data stored in the hard drive.

3.13 DIELECTRICS

Dielectric is basically an insulator. If the main function of an insulator is to provide insulation, it is called an *insulator*. If the main function of the insulator is to store electric charges, it is called *dielectric*. The dielectric is used in capacitor where the charge is stored. The dielectric is characterized by (i) dielectric constant (ii) dielectric loss (iii) dielectric strength and (iv) resistivity.

Here, the dielectric constant and dielectric loss are strongly frequency dependent. Dielectric loss manifests itself as heat and is used in food processing (Microwave cooking).

When a dielectric is placed in an electric field, it undergoes some changes and hence, electric charges are stored. In the process of storing charges, one side of the dielectric becomes positive and the other side becomes negative. Hence, the name dielectric.

From the energy band diagram of a dielectric shown in Fig. 3.47, it is clear that the electrons are bound to the nucleus and there is no free electron.

The forbidden gap E_G shown in Fig. 3.47 is very large in dielectrics and excitation of electrons from the normally full valence band to the empty conduction band is not possible under ordinary conditions. Therefore, conduction cannot occur in a dielectric. Even if a dielectric contains impurities, extrinsic conduction is not possible. The resistivity of an ideal dielectric should be high. However, in practice, dielectrics conduct electric current which is negligible and their resistivities range from 10^{10} to $10^{20} \Omega \cdot m$.



Fig. 3.47 Energy band diagram of a dielectric

3.14 DIELECTRIC CONSTANT

The electric properties of dielectric materials are usually described in terms of the dielectric constant. Dielectric constant of a dielectric is defined and measured as the ratio of capacitance of a capacitor containing the dielectric medium to the capacitance of the same capacitor with air as the medium. Therefore,

$$\varepsilon_r = \frac{C}{C_0}$$

where C_0 is the capacitance with air as the medium between the plates and *C* is the capacitance with dielectric as medium. Here, the dielectric constant or relative permittivity of dielectric, ε_r , is a dimensionless quantity. It is always greater than unity and is independent of the size or shape of the dielectric. In fact, ε_r describes the ability of the dielectric material to store electric charges. At the same time, another quantity known as the absolute permittivity of the medium, ε_r , is used. It is given by $\varepsilon = \varepsilon_0 \varepsilon_r$, where $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m represents the permittivity of free space.



Fig. 3.48 Parallel plate capacitor (a) without dielectric (b) with dielectric

Consider the two plates A and B of a parallel plate capacitor connected to a DC voltage source V_0 without dielectric as shown in Fig. 3.48(a). The potential difference across the two plates is V_0 . The charges on the two plates are $+q_0$ and $-q_0$ when the potential difference is V_0 . The capacitance of the capacitor without dielectric i.e. with air or free space as the medium is

$$C_0 = \frac{q_0}{V_0}$$

When the voltage source V_0 is removed, the circuit is open. There is no path for the energy to flow or charge flow out or in. Hence q_0 (number of charges) on either plate is a constant. Now a dielectric is inserted in the form of a rectangular slab in between the plates as shown in Fig. 3.48(b). The potential decreases to a value V which is measured using a voltmeter. If the slab is removed, $V \rightarrow V_0$. This means that the original charges on the plates are not affected. They remain as $+q_0$ and $-q_0$. This means that the capacitance of the condenser (storing charges/unit voltage) has increased, say from C_0 to C.

Then,

Let

 $C = \frac{q_0}{V}$ $V = \frac{V_0}{\varepsilon_r}, \text{ where } \varepsilon_r > 1 \text{ as } V < V_0$

Then,

$$C = \frac{q_0}{V} = \frac{q_0}{V_0 / \varepsilon_r} = \frac{q_0}{V_0} \varepsilon_r = \varepsilon_r C_0$$

$$\varepsilon_r = \frac{C}{C_0}$$
, which is greater than 1.

Also for a parallel plate capacitor, $C_0 = \frac{\varepsilon_0 A}{d}$ and $C = \frac{\varepsilon_r \varepsilon_0 A}{d}$

i.e., $C > C_0$ with a dielectric as $\varepsilon_r > 1$.

To understand this, a process called polarization has to be studied.

3.15 POLARIZATION – AN ATOMIC VIEW

The effect on an atom in a dielectric in the absence and presence of a field is discussed in this section. In the

absence of a field, the centers of positive and negative charges coincide as shown in Fig. 3.49(a) and when a field is applied from right to left, the charges are displaced as shown in Fig. 3.49(b).

It is observed that the centers of the positive and negative charges no longer coincide. There is a charge separation that has been induced i.e. a dipole is induced. This process of producing induced dipoles oriented along the field direction is known as *polarization*.



By making use of this polarization process, the reason for decrease in voltage and increase in capacitance with a dielectric can be demonstrated as follows.

3.16 DIELECTRIC POLARIZATION

Consider a slab of dielectric located between the plates of a parallel plate capacitor. In the absence of an external electric field, each elementary volume of the dielectric has no dipole moment. Hence, the polarization

is zero. When the electric field is switched ON, dipoles are induced, which form chains along the field lines, as shown in Fig. 3.50.

Thus, the action of electric field brings the dipoles into a certain ordered arrangement in space. It is seen that the ends of adjacent dipoles carry opposite charges and neutralize each other. Only the charges of the dipole ends terminating on the opposite faces of the slab remain uncompensated. Thus, as seen from Fig. 3.50, dipoles are induced due to the action of the field in the dielectric. This is known as *dielectric polarization*.

Let E_0 be the field and V_0 the potential difference without dielectric for a parallel plate capacitor. When the medium between the plates is air, then

$$E_0 = \frac{V_0}{d}$$



Fig. 3.50 Polarization of a dielectric

When the dipoles are induced, all the charges except the negative on the left (side P) and the positive in the right (side Q) cancel each other. What is left out is, negative charges of surface charge density $+\sigma_p$ at Q and $-\sigma_p$ at P (Bound charges) creating a field E_i from Q to P opposite to the direction of E_0 . Hence, the effective field is reduced. As seen earlier, q_0 does not change. Therefore,

$$\overline{E} = (\overline{E}_0 - \overline{E}_i) < \overline{E}_0$$
, (as 'd' is same and $V < V_0$)

Since |E| = V / d, we get

 $V = |E| d = (E_0 - E_i) d < V_0$

Thus, the effect of polarization is to reduce the magnitude of the external field \overline{E}_0 . The induced surface charges on the dielectric give rise to an induced electric field \overline{E}_i which opposes the external field \overline{E}_0 and weakens it. Therefore, the net electric field \overline{E} in the dielectric has a magnitude given by

$$E = E_0 - E_i$$

Therefore, $C = \frac{q_0}{V}$ i.e., $C_0 = \frac{q_0}{V_0}$

Therefore, the capacity of a dielectric depends on the number of dipoles induced per unit volume for unit field. This requires a knowledge of the polarizability α , polarization vector \overline{P} and electric flux density or electric displacement vector \overline{D} .

3.16.1 Polarizability, α

If p_{ind} is the induced dipole moment, then

$$p_{ind} \propto E$$
 or $p_{ind} = \alpha E$

As the electric field increases, the strength of induced dipole also increases. Hence, α is called the polarizability.

3.16.2 Polarization Vector, \overline{P}

The polarization vector \overline{P} is defined as the dipole moment (*p*) per unit volume. From Fig. 3.51, the induced charges (*q*⁴) per unit area can be represented by

$$P = \frac{p}{\text{volume}} = \frac{q'l}{Al} = \frac{q'}{A} = \sigma_p,$$

(since dipole moment, p = q'l)

where P is the magnitude of polarization vector \overline{P} and σ_p is the surface polarization charge density or induced charge density.

Thus, the polarization vector is also defined as induced charge per unit area with the polarization vector \overline{P} normal to the area. The vector \overline{P} is always directed from negative to positive.

If p is the average dipole moment per unit volume and N is the number of moles per volume, then the magnitude of polarization vector \overline{P} is

$$P = Np \tag{3.13}$$



Fig. 3.51 Formation of dipole with direction of polarization

It follows that polarization is equal to the surface density of the induced charges in a dielectric.

3.16.3 Electric Flux Density (or) Displacement Vector \overline{D}

To arrive at an explanation for \overline{D} , Gauss law must be applied to a dielectric.



Fig. 3.52 Gauss law applied to a parallel plate capacitor (a) without dielectric (b) with dielectric

Consider the application of Gauss law to a parallel plate capacitor without a dielectric first and then with a dielectric. Figure 3.52(a) shows a Gaussian surface drawn enclosing the charge q on one of the plates. Here, q is known as the free charge on the capacitor plate. The electric field \overline{E}_0 at any point on the Gaussian surface is given by

$$\oint \overline{E}_0 \cdot \overline{dA} = \frac{q}{\varepsilon_0} \text{ or } E_0 A = \frac{q}{\varepsilon_0}$$

$$E_0 = \frac{q}{\varepsilon_0 A}$$

$$E_0 = \frac{\sigma}{\varepsilon_0}$$
(3.14)

i.e.,

or

$$E_0$$

where $\sigma = q/A$ is the surface charge density of free charges on plates A and B in coulomb/m².

Consider the capacitor with a dielectric. Due to polarization, charges are induced on the surfaces P and Q of the dielectric resulting in a field E_i opposite to E_0 as shown in Fig. 3.50. Now a Gaussian surface is drawn as shown in Fig. 3.52(b) enclosing the free charges on the capacitor plate and induced charges on the dielectric surface. Let q' be the induced charge on the surface of the dielectric which is called the bound charge.

(3.16)

Note that q' is negative charge. Then (q - q') is the net charge within the Gaussian surface. Let \overline{E} be the resultant field inside the dielectric. Then, according to Gauss theorem

$$\oint \overline{E} \cdot \overline{dA} = \frac{q - q'}{\varepsilon_0} \tag{3.15}$$

or

or

$$E = \frac{\sigma}{\varepsilon_0} - \frac{\sigma_p}{\varepsilon_0}$$
 where $\frac{q}{A} = \sigma$ and $\frac{q'}{A} = \sigma_p$

Here, $\sigma_p = P$, the surface density of induced charges and σ is the surface density of free charges. The magnitude of the induced field can be written as

$$E_i = \frac{\sigma_p}{\varepsilon_0}$$

 $E = \frac{q}{\varepsilon_0 A} - \frac{q'}{\varepsilon_0 A}$

Therefore, the magnitude of the effective field is

$$E = E_0 - E_i$$

Equation (3.16) indicates that the induced surface charge density σ_p weakens the original field when the dielectric is present.

Since
$$\frac{C}{C_0} = \varepsilon_r$$
,
 $\frac{C}{C_0} = \frac{q_0/V}{q_0/V_0} = \frac{V_0}{V} = \frac{E_0}{\overline{E}}$
Hence,
 $\frac{E_0}{E} = \varepsilon_r$

Thus, the initial field and the resultant field are related through the relation,

$$E_0 = \varepsilon_r E$$

Substituting the value of E from Eqn. (3.16), we get

$$E = \frac{E_0}{\varepsilon_r} = \frac{q}{\varepsilon_0 A} - \frac{q'}{\varepsilon_0 A}$$

Using Eqn. (3.14), we can write the above relation as

$$\frac{q}{\varepsilon_0 \varepsilon_r A} = \frac{q}{\varepsilon_0 A} - \frac{q'}{\varepsilon_0 A}$$
$$\frac{q}{\varepsilon_r} = q - q'$$
(3.17)

Using Eqn. (3.17) in Eqn. (3.15), we get

$$\oint \overline{E} \cdot \overline{dA} = \frac{q - q'}{\varepsilon_0} = \frac{q}{\varepsilon_0 \varepsilon_r}$$

or
$$\varepsilon_0 \varepsilon_r \int \overline{E} \cdot \overline{dA} =$$

 $\varepsilon(EA) = q$ (since $e = \varepsilon_0 \varepsilon_r$) or $E = \frac{\sigma}{c}$ or $\sigma = \varepsilon E$

i.e.,

Three Field Vectors 3.16.4

The resultant magnitude of electric field inside the dielectric is given by Eqn. (3.16) as

$$E = \frac{q}{\varepsilon_0 A} - \frac{q'}{\varepsilon_0 A}$$

We know that, $\frac{q'}{A} = \sigma_p = P$. Therefore,

$$\frac{q}{\varepsilon_0 A} = E + \frac{P}{\varepsilon_0}$$

or

 $\frac{q}{\Lambda} = \varepsilon_0 E + P$ The quantity q/A is called electric displacement, D.

Thus, $D = \frac{q}{A} = \sigma = \varepsilon E$ is the surface density of free charges.

Hence, the electric displacement can be written in vector form as

$$\overline{D} = \varepsilon_0 \overline{E} + \overline{P} \tag{3.19}$$

(3.18)

The above equation can be written as

$$\overline{P} = \overline{D} - \varepsilon_0 \overline{E} = \varepsilon \overline{E} - \varepsilon_0 \overline{E} \qquad (\text{since } \overline{D} = \varepsilon \overline{E})$$
$$= \varepsilon_0 \varepsilon_r \overline{E} - \varepsilon_0 \overline{E} \qquad (\text{since } \varepsilon = \varepsilon_0 \varepsilon_r)$$
$$= \varepsilon_0 (\varepsilon_r - 1) \overline{E}$$
$$\overline{P} = \chi_e \varepsilon_0 \overline{E}$$

where $\chi_e = (\varepsilon_r - 1)$ is called the electrical susceptibility. Here, the relative permittivity ε_r of the dielectric material can be written as

$$\varepsilon_r = 1 + \chi_e$$

Therefore, $\overline{P} \propto \overline{E}$ and measures the additional flux density arising from the presence of dielectric material as compared to free space.

The three electric vectors, \overline{E} , \overline{D} and \overline{P} are shown in Fig. 3.53. The expression for \overline{D} shows that this vector is related to the free charges only; they are the charges stored on the capacitor plates. The expression for \overline{P} shows that it is related to the bound charges only. The expression for the vector \overline{E} shows that it is connected with both types of charges present. It may be seen from Fig. 3.53 that the lines of \overline{D} begin and end on free charges; the lines of P begin and terminate on induced charges and the lines of E change from one medium to the other.


Fig. 3.53 Field vectors E, D and P

EXAMPLE 3.10

Calculate the polarization produced in dielectric medium of dielectric constant 6 when it is subjected to an electric field of 100 V/m. $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m).

Solution Given $\varepsilon_r = 6$, $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m and E = 100 V/m.

Polarization, $P = \varepsilon_0 (\varepsilon_r - 1)E = 8.854 \times 10^{-12} \times (6 - 1) \times 100 = 4.425 \times 10^{-9} \text{ C/m}^2$

EXAMPLE 3.11

When NaCl crystal is subjected to an electric field of 50 V/cm, the resulting polarization is 2.215×10^{-7} C/m². Calculate relative permittivity of NaCl.

Solution Given E = 50 V/cm and $P = 2.215 \times 10^{-7}$ C/m².

The magnitude of polarization is

$$P = \varepsilon_0 (\varepsilon_r - 1)E$$

Hence, the relative permittivity is

$$\varepsilon_r = 1 + \frac{P}{\varepsilon_0 E} = \frac{2.215 \times 10^{-7}}{8.854 \times 10^{-12} \times 50 \times 100} = 6.006$$

EXAMPLE 3.12

The capacitance of a parallel plate capacitor is 2 μ F. The dielectric has permittivity, $\varepsilon_r = 80$. Find the energy stored in the capacitor for an applied voltage of 1 KV.

Solution Given $C = 2 \mu F = 2 \times 10^{-6} F$, $\varepsilon_r = 80$ and V = 1 KV = 1000 V.

Energy stored in a capacitor, $E = \frac{1}{2}CV^2 = \frac{1}{2} \times 2 \times 10^{-6} \times (1000)^2 = 1 \text{ J}$

EXAMPLE 3.13

Find the capacitance of a layer of Al_2O_3 that is 0.5 μ m thick and 2000 mm² of square area $(\varepsilon_0 = 8.854 \times 10^{-12} \text{ F/m})$

Solution Given $A = 2000 \text{ mm}^2$, $d = 0.5 \,\mu\text{m}$ and $\varepsilon_0 = 8.854 \times 10^{-12} \text{ F/m}$.

Capacitance,

$$C = \frac{\varepsilon_0 \varepsilon_r A}{d} = \frac{8.854 \times 10^{-12} \times 2000 \times 10^{-6}}{0.5 \times 10^{-6}} = 3.54 \times 10^{-8} = 0.354 \,\mathrm{nF}$$

EXAMPLE 3.14

A layer of porcelain is 80mm long, 20 mm wide and 0.7 μ m thick. Calculate its capacitance with $\varepsilon_r = 6$.

Solution Given Area = $l \times b = 80 \times 20 \times 10^{-6}$ m, $d = 0.7 \,\mu$ m and $\varepsilon_r = 6$.

Capacitance,

$$C = \frac{\varepsilon_0 \varepsilon_r A}{d} = \frac{8.854 \times 10^{-12} \times 6 \times 80 \times 20 \times 10^{-6}}{0.7 \times 10^{-6}} = 1.21 \times 10^{-7} = 0.121 \mu \text{F}$$

POLARIZATION PROCESSES – TYPES OF POLARIZATION 3.17

When a specimen is placed in a DC electric field, four types of polarization processes can occur, namely (i) Electronic polarization, (ii) Ionic polarization, (iii) Orientational polarization and (iv) Space charge polarization.

Electronic Polarization

Electronic polarization occurs due to the displacement of positively charged nucleus and negatively charged electrons in opposite directions, when an external electric field is applied, and thereby, an induced dipole moment is created in the dielectric. Monoatomic gases exhibit this kind of polarization. Materials in which the dipoles are only induced by an external field are called non-polar and often referred to as non-polar dielectrics. The electronic polarization sets in 10^{-14} to 10^{-15} secs.

Since the charge separation is proportional to electric field E, the electronic dipole moment is also proportional to electric field.

i.e.,

$$p_e \propto E$$

 $p_e = \alpha_e$

or

$$\alpha_e E$$

(3.20)

R

where α_e is the electronic polarizability and p_e is the electronic dipole moment.

Derivation of an expression for electronic polarizability, α_e Consider an atom of atomic number Z. Then the nuclear charge is +Ze and the electron cloud has a charge -Ze in a sphere of radius R as shown in Fig. 3.54.

Hence, the charge density is given by

$$\rho = \frac{\text{charge}}{\text{volume}} = \frac{-Ze}{\frac{4}{3}\pi R^3}$$

$$\rho = \frac{-3}{4} \left(\frac{Ze}{\pi R^3}\right)$$
(3.21)
Fig. 3.54 Sphere of radius R with nuclear charge and electron cloud

i.e.,

When an external field is applied, the nucleus and the electrons experience a Lorentz force with magnitude ZeE acting in opposite directions and hence, are pulled apart by a distance x. This can be treated as an electron cloud shifting by a certain distance as shown in

Fig. 3.55. Due to this separation, a coulomb force which opposes the Lorentz force is induced.

The Lorentz force = -ZeE (3.22)

Coloumb force = $\frac{Ze}{4\pi\varepsilon_o x^2}$ × charge enclosed in a sphere of radius *x*.

Hence, Coulomb force =
$$Ze \times \frac{\frac{4}{3}\pi x^3 \rho}{4\pi\epsilon_0 x^2}$$

Substituting Eqn. (3.21) in the above equation, we get

Coulomb force =
$$Ze\left(\frac{\frac{4}{3}\pi x^3 \left\{\frac{-3}{4}\left(\frac{Ze}{\pi R^3}\right)\right\}}{4\pi \varepsilon_0 x^2}\right)$$

= $Ze\left(\frac{-Zex}{4\pi \varepsilon_0 R^3}\right)$

i.e., Coulomb force = $\left(\frac{-Z^2 e^2 x}{4\pi\varepsilon_0 R^3}\right)$

At equilibrium, Lorentz force = Coulomb force.

$$\left(\frac{-Z^2 e^2 x}{4\pi\varepsilon_0 R^3}\right) = -ZeE$$

Ze

i.e.,

We know that the electronic dipole moment is the product of the charge and distance between the charges.

i.e.,
$$p_e = Zex$$
 (3.24)

Substituting the value of x in Eqn. (3.24), we obtain

$$p_e = 4\pi\varepsilon_0 R^3 E \tag{3.25}$$

Using Eqns. (3.20) and (3.25), we get

$$\alpha_e = \frac{p_e}{E} = 4\pi\varepsilon_0 R^3 \tag{3.26}$$



Fig. 3.55 Shifting of electron cloud by distance x

(3.23)

From the above equation, it is clear that, the electronic polarizability α_e is independent of temperature and proportional to the volume of the atoms.

Using Eqn. (3.13), the magnitude of electronic polarization vector \overline{P}_e can be written as

$$P_e = Np_e = N\alpha_e E$$
 [since $p_e = \alpha_e E$ from Eqn. (3.20)]

In vector form, $\overline{P}_e = N\alpha_e\overline{E}$

where N is the number of atoms or moles per unit volume.

From section 3.16.4, we have

$$P_e = \varepsilon_0 (\varepsilon_r - 1)E$$

Comparing the above two equations, we get

$$\varepsilon_0(\varepsilon_r - 1)E = N\alpha_e E$$

$$\alpha_e = \frac{\varepsilon_0(\varepsilon_r - 1)}{N}$$
(3.27)

i.e.,

Also, we know that $\varepsilon_r = 1 + \chi_e$

Substituting $\chi_e = \frac{P_e}{\varepsilon_0 E}$, we get

$$\varepsilon_r = 1 + \frac{P_e}{\varepsilon_0 \overline{E}}$$

Substituting $P_e = N\alpha_e E$ in the above equation, we get

$$\varepsilon_r = 1 + \frac{N\alpha_e}{\varepsilon_0} \tag{3.28}$$

The above expression indicates that the contribution to dielectric constant ε_r due to electronic polarization alone is true for a non-polar gas. Substituting Eqn. (3.26) in the above equation, we obtain the value of relative permittivity for a monoatomic gas as

$$\varepsilon_r = 1 + 4\pi NR^3$$

EXAMPLE 3.15

Calculate the electronic polarizability of neon. The radius of neon atoms is 0.158 nm. $(\varepsilon_0 = 8.854 \times 10^{-12} \text{ F/m}).$

Solution Given R = 0.158 nm and $\varepsilon_0 = 8.854 \times 10^{-15}$ F/m.

Electronic polarizability, $\alpha_e = 4\pi\epsilon_0 R^3 = [4 \times 3.14 \times 8.85 \times 10^{-12} \times (0.158 \times 10^{-9})^3]$

$$= 4.386 \times 10^{-40} \,\mathrm{F} \cdot \mathrm{m}^2$$

EXAMPLE 3.16

The dielectric constant of a helium gas at NTP is 1.0000684. Calculate the electron polarizability of helium atoms if the gas contain 2.7×10^{26} atoms/m³ and hence, calculate the radius of helium atom. ($\varepsilon_0 = 8.854 \times 10^{-12}$ F/m)

Solution Given $\varepsilon_r = 1.0000684$, $N = 2.7 \times 10^{26}$ atoms/m³ and $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m.

(i) Electron polarizability,
$$\alpha_e = \frac{\varepsilon_0(\varepsilon_r - 1)}{N}$$

 $= \frac{8.854 \times 10^{-12} \times (1.0000684 - 1)}{2.7 \times 10^{26}} = 2.242 \times 10^{-42} \,\mathrm{F \cdot m^2}$
(ii) Radius of helium atom, $R = \left[\frac{\alpha_e}{4\pi\varepsilon_0}\right]^{1/3}$ (since $\alpha_e = 4\pi\varepsilon_0 R^3$)
 $= \left[\frac{2.24 \times 10^{-42}}{4 \times 3.14 \times 8.854 \times 10^{-12}}\right]^{1/3} = 0.27 \times 10^{-10} = 0.27 \,\mathrm{\AA}$

EXAMPLE 3.17

Calculate the electronic polarizability of Argon atom, if $\varepsilon_r = 1.0024$ at NTP and $N = 2.7 \times 10^{25}$ atoms/m³.

Solution Given $\varepsilon_r = 1.0024$ and $N = 2.7 \times 10^{25}$ atoms/m³.

Relative permittivity, $\varepsilon_r = 1 + \frac{N\alpha_e}{\varepsilon_0}$

Therefore,
$$\alpha_e = \varepsilon_0 \frac{(\varepsilon_r - 1)}{N} = \frac{(8.854 \times 10^{-12}) \times (1.0024 - 1)}{2.7 \times 10^{25}} = 7.9 \times 10^{-40} \,\mathrm{F \cdot m^2}$$

EXAMPLE 3.18

The number of atoms in hydrogen gas is 9.8×10^{20} atoms/cc. The radius of hydrogen atom is 0.053 nm. Calculate its electronic polarizability and relative permittivity.

Solution Given $R = 0.053 \times 10^{-9}$ m and $N = 9.8 \times 10^{20}$ atoms/cc.

Electronic polarizability, $\alpha_e = 4\pi\varepsilon_0 R^3 = [4 \times 3.14 \times 8.85 \times 10^{-12} \times (0.053 \times 10^{-19})^3]$

$$= 1.657 \times 10^{-41} \,\mathrm{F} \cdot \mathrm{m}$$

Relative permittivity, $\varepsilon_r = 1 + 4\pi NR^3 = 1 + 4 \times 3.14 \times 9.8 \times 10^{26} \times (0.053 \times 10^{-9})^3 = 1.0018$

Ionic polarization

This type of polarization occurs in ionic crystals. In ionic crystals, positive ions are called *cations* and negative ions are called *anions*. The polarization that occurs due to displacement of cations and anions when an electric field *E* is applied is called ionic polarization. When the direction of field applied is from left to right,

the cations move to the right and the anions move to the left. i.e., they are displaced from their equilibrium positions as in Fig. 3.56(b). The ionic polarization takes 10^{11} to 10^{14} secs to set in.

	Applied field, E	
No field		
$\ominus \oplus \ominus \oplus \ominus$	\ominus \oplus \ominus \oplus \ominus	
$\oplus \ \ominus \ \oplus \ \ominus \ \oplus$	$\oplus \ominus \oplus - \oplus$	
$\ominus \oplus \ominus \oplus \ominus$	$\bigcirc \oplus \ \bigcirc \oplus \ \bigcirc$	
$\oplus \ \ominus \ \oplus \ \ominus \ \oplus$	$ \oplus \bigcirc \oplus \bigcirc \oplus $	
(a)	(b)	

Fig. 3.56 Ionic polarization (a) unpolarized dielectric and (b) ionic polarization

The dipole moment induced in ionic crystals is proportional to the applied electric field and is expressed as

$$p_i = \alpha_i E \tag{3.29}$$

where α_i is the ionic polarizability and p_i is the ionic dipole moment.

Derivation of α_i Let us assume an ionic crystal with one cation and one anion as shown in Fig. 3.57(a). When an electric field E is applied, the cations move to the right by a distance, x_1 and the anions move to the left by a distance x_2 as shown in Fig. 3.57(b).

Therefore, the resultant dipole moment is given by

$$p_i = e(x_1 + x_2)$$



Let the restoring force constants of the cations and anions be β_1 and β_2 respectively.

If F is the magnitude of force due to the applied electric field, then

$$F = \beta_1 x_1 = \beta_2 x_2$$
$$x_1 = \frac{F}{\beta_1} \text{ and } x_2 = \frac{F}{\beta_1}$$

i.e.,

Here, β_1 and β_2 depend on the mass of the ions and ω_0 the angular frequency of the molecules in which the ions are present. Then,

$$F = eE$$
, $\beta_1 = m\omega_0^2$ and $\beta_2 = M\omega_0^2$,

Therefore,

where *M* is the mass of the negative ions and *m* is the mass of the positive ions.

 $x_1 = \frac{eE}{m\omega_0^2}$ and $x_2 = \frac{eE}{M\omega_0^2}$

Substituting x_1 and x_2 in Eqn.(3.30), we get

$$p_i = \frac{e^2 E}{\omega_0^2} \left[\frac{1}{M} + \frac{1}{m} \right]$$
(3.31)



(3.30)

and (b) with electric field

Using Eqn.(3.29), we get

$$\alpha_i = \frac{p_i}{E}$$

Substituting Eqn. (3.31) in the above equation, we get

$$\alpha_i = \frac{e^2}{\omega_0^2} \left[\frac{1}{M} + \frac{1}{m} \right]$$
(3.32)

i.e.,

 $\alpha_i = \frac{e^2}{\omega_0^2} \left\lfloor \frac{M+m}{Mm} \right\rfloor$

Therefore, the ionic polarizability is given by

$$\alpha_i = \frac{e^2}{m_r \omega_0^2} \tag{3.33}$$

where $m_r = \frac{Mm}{M+m}$ is the reduced mass.

From Eqn. (3.33), it is clear that α_i does not depend on temperature.

In addition, the ions experience electronic polarization since an electric field is applied.

The ionic polarization in vector form is given by

$$P_i = N\alpha_i E$$

Substituting Eqn. (3.32) in the above equation, we get

$$\overline{P}_i = N \left[\frac{e^2}{\omega_0^2} \left(\frac{1}{M} + \frac{1}{m} \right) \right] \overline{E}$$

Orientational Polarization

The orientational polarization occurs in polar dielectrics, which consist of molecules with permanent dipole moment. Any molecule has two charges i.e., positive and negative charges. In some molecules, called polar molecules, the center of gravity of these charges are separated even in the absence of an external electric field. Hence, a polar molecule is said to have an intrinsic dipole moment and carries a permanent dipole. When a polar molecule is placed in a uniform electric field \overline{E} , the field exerts a force $+q\overline{E}$ on the positive charge



Fig. 3.58 Force on a dipole

and $-q\overline{E}$ on the negative charge. The net force acting on the dipole is zero since the two forces acting on it are equal and opposite to each other. Therefore, there is no translational force acting on the dipole in a uniform electric field. However, the forces are anti-parallel and constitute a couple which tends to rotate the dipole as shown in Fig. 3.58. Therefore, the magnitude of torque acting on the dipole is given by

$$\tau = (Force)(Distance) = qEd \sin \theta$$

where q is the magnitude of the charge and d is the distance between the positive and negative charges. But the magnitude of dipole moment is p = qd.

Therefore, $\tau = pE\sin\theta$

In vector form, $\overline{\tau} = \overline{p} \times \overline{E}$

Thus, a dipole experiences a torque in a uniform electric field and hence rotates so that it aligns with the direction of applied electric field. But, a free dipole aligns its axis in the direction of field.

Further, the electric field induces a dipole moment in the molecule. Therefore, the total dipole moment of the molecule is the sum of the induced and permanent dipole moments.

i.e., $p_T = p_{ind} + p_{per}$

where p_T , p_{ind} and p_{per} are the total, induced and permanent dipole moments respectively.

For polar molecules, $p_{\text{ind}} \ll p_{\text{per}}$, we get, $p_T = p_{\text{per}}$.

In the absence of external electric field, the orientation of dipoles is random resulting in a complete cancellation of each other's effect as shown in Fig. 3.59(a). In the presence of an electric field, the dipoles rotate about their axis of symmetry to align with the applied field. In orientation polarization, restoring forces do not exist.



Fig. 3.59 Molecular dipoles in the (a) absence of electric field and (b) presence of electric field

Orientational polarization is strongly temperature dependent. At high temperatures, due to thermal agitation, the dipoles turn through a small angle, as shown in Fig. 3.59(b). Even in case of liquids or gases, where molecules are free to rotate, a complete alignment cannot be achieved due to the randomizing effect of the temperature. However, it is proved that it is enough if one molecular dipole completely aligns with the direction of the applied field so that the order of orientation and electronic polarization are same. In case of solids, the molecules are fixed in their positions and their rotation is highly restricted by the lattice forces, leading to a great reduction in their contribution to orientation polarization. Due to this, the dielectric constant of water is 80 and solid ice is 10.

Since the orientation polarization involves rotation of molecules, it takes relatively longer time when compared to electronic and ionic polarizations. The buildup time of orientation polarization is of the order of 10^{-10} s or more.

For example, consider the structures of CH_4 (Methane) and $CHCl_3$ (Chlorofom) which are shown in Fig. 3.60 (a) and (b). Since the centres of +ve and –ve charges coincide in CH_4 , there is no permanent dipole moment. But in $CHCl_3$, the centres of +ve and –ve charges do not coincide.



Therefore, there exists a permanent dipole moment. Since chlorine is more electronegative than hydrogen, it pulls the electrons to itself and hence, dipoles are created. These dipoles are oriented in a random manner. When an electric field is applied, the dipoles tend to align in the direction of the electric field.

From Langevin's theory, $\overline{P}_o = N \overline{p}_{avg} = \frac{N p_o^2 \overline{E}}{3kT} = N \alpha_o \overline{E}$

where $\overline{P_o}$ is the orientational polarization vector, $\overline{P_{avg}}$ is the average dipole moment due to orientational polarization, k is the Boltzmann's constant and T is the absolute temperature.

Therefore, the orientation polarizability α_0 is given by

$$\alpha_o = \frac{p_o^2}{3kT}$$

where p_0 is the orientational dipole moment.

Space Charge Polarization

Space charge or interfacial or migrational, \overline{P}_m polarization occurs in heterogeneous dielectric materials in which there is a change of electrical properties between different phases. Also, it occurs in homogeneous dielectrics containing impurities, pores filled with air, inclusions of hygroscopic water etc. In particular, the properties of homogeneous materials in their outer layers are likely to differ because of the environmental effects and as a result, they behave like heterogeneous materials. When an electric field is applied, the electric charges migrate within the impurity regions and store up at the interfaces. The accumulation of charges takes place with opposite polarity on the interfaces as shown in Fig. 3.61. The space charge polarization takes generally a longer time of 10^{-2} secs. This polarization occurs at low (power) frequencies.



Fig. 3.61 Space Charge Polarization

Interfacial polarization occurs whenever there is an accumulation of charge at an interface between two materials or between two regions within a material as shown in Fig. 3.61(a). One of the typical examples of interfacial polarization is the grain boundaries. It can trap charges migrating under the influence of an

applied electric field as shown in Fig. 3.61(b). Interfacial polarization also arises in heterogeneous dielectric materials when there is a dispersed phase within a continuous phase as shown in Fig. 3.61(c).

Normally, interfacial polarization exists in all materials since it passes crystal defects, impurities, and various mobile charge carriers such as electrons, holes or ionized impurity ions.

Since there is no general procedure to determine the contribution of the changes to the total polarization of a dielectric, interfacial polarization is often omitted. However, it is important in practical aspects.

Total Polarization

The total polarization in a material \overline{P}_{total} in vector form is given by

$$\overline{P}_{\text{total}} = \overline{P}_e + \overline{P}_i + \overline{P}_o + \overline{P}_m$$

Since interfacial polarization or migrational polarization is omitted, the magnitude of total polarization in a material is given by

$$\overline{P}_{\text{total}} = \overline{P}_e + \overline{P}_i + \overline{P}_o$$

Hence, the total polarization of the material is

$$P_{\text{total}} = N[\alpha_e + \alpha_i + \alpha_o] E$$

The total polarizability is given by

$$\alpha = \alpha_e + \alpha_i + \alpha_0$$

Substituting the values of α_e , α_i , and α_0 from the previous discussion, we get

$$\alpha = 4\pi\varepsilon_0 R^3 + \frac{e^2}{\omega_0^2} \left(\frac{1}{M} + \frac{1}{m}\right) + \frac{p_o^2}{3kT}$$

$$\overline{P} = N \left[4\pi\varepsilon_0 R^3 + \frac{e^2}{\omega_0^2} \left(\frac{1}{M} + \frac{1}{m}\right) + \frac{p_o^2}{3kT}\right] \overline{E}$$
(3.34)

Therefore,

The above equation is called the Langevin-Debye Equation.

It is not necessary that all types of polarization must exist in a dielectric. It is possible for one or more of the contributions to the polarization to be either absent or negligible in magnitude relative to the others.

3.18 LOCAL OR INTERNAL FIELD

When a dielectric material is placed inside an electric field, due to polarization of the dipoles, an internal field arises. In case of gases, the separation between atoms is large and hence, interaction between atoms can be neglected. When an external field \overline{E} is applied, the internal field experienced by an atom in the gas is also \overline{E} . But in solids and liquids, the atoms are close to each other and there is strong interaction between atoms. Hence, the internal field at any given point inside will be greater than the external field \overline{E} . Let \overline{E}_{int} or \overline{E}_{local} be the internal field acting at the location of the atom as given by

 \overline{E}_{int} = Applied field + sum of field due to neighbouring atoms

Lorentz Method for Finding Internal Field in Cubic Structures

Consider a dielectric placed between the plates *X* and *Y* of a parallel plate capacitor and energized using an electric field as shown in Fig. 3.62. Let σ be the surface density due to free charges on the plates *X* and *Y*.

Consider a spherical cavity around an atom A inside the dielectric. The radius of the spherical cavity is larger when compared to the radius of the atom. The internal field \overline{E}_{int} at A in vector form is given by

(3.35)

$$\overline{E}_{\text{int}} = \overline{E}_1 + \overline{E}_2 + \overline{E}_3 + \overline{E}_4$$

where \overline{E}_1 is the field intensity at A due to surface density of free charges on plates X and Y and is given by

$$\overline{E}_1 = \frac{D}{\varepsilon_0} \tag{3.36}$$

From Eqn. (3.19), we have $\overline{D} = \varepsilon_0 \overline{E} + \overline{P}$.

Therefore,

$$\overline{E}_1 = \frac{\varepsilon_0 \overline{E} + P}{\varepsilon_0} = \overline{E} + \frac{P}{\varepsilon_0}$$

The field \overline{E}_2 is the field intensity at A due to induced surface charge density on the two sides of the dielectric.

$$\overline{E}_2 = \frac{-\overline{P}}{\varepsilon_0}$$
 (since \overline{E}_2 and \overline{P} are opposite to each other)

The field \overline{E}_3 is the field intensity at A due to other atoms in the cavity.

For cubic structures or spherical symmetry, $\overline{E}_3 = 0$.

The field \overline{E}_4 is the field intensity due to polarization charges on the surface of the cavity.

To determine \overline{E}_4 : The enlarged view of the spherical surface is shown in Fig. 3.63.



Fig. 3.63 Calculation of \overline{E}_4 using enlarged view of the spherical surface

In the sphere, consider an elementary ring of an area dA lying between angles θ and $\theta + d\theta$ with reference to the direction of the applied field. The radius of the sphere is *r*.

Therefore, ring area $dA = (2\pi y) \times$ thickness of the ring



Fig. 3.62 Schematic diagram for Lorentz internal field calculation

$$=(2\pi y)\times ds=2\pi yrd\theta\times ds$$

But $\sin \theta = y/r$

i.e., $y = r\sin\theta$

and
$$dA = 2\pi r^2 \sin \theta \, d\theta$$

But the component normal to \overline{ds} is \overline{P}_N .

i.e.,
$$\overline{P}_N = \overline{P} \cos \theta$$

Thus, the charges on *ds* will be

$$dq = P_N \cdot ds$$

$$dq = P\cos\theta\,ds$$



Fig. 3.64 Normal component of polarisation

The differential magnitude of electric field at A due to the differential charge dq along AB is

$$dE_4 = \frac{dq}{4\pi\varepsilon_0 r^2} = \frac{P\cos\theta \, ds}{4\pi\varepsilon_0 r^2}$$

The differential component of this intensity along the field direction E is

$$dE_4 = \left(\frac{P\cos\theta \, ds}{4\pi\varepsilon_0 r^2}\right)\cos\theta \tag{3.37}$$

and the perpendicular component of the field is

$$dE_4 = \left(\frac{P\cos\theta\,ds}{4\pi\varepsilon_0 r^2}\right)\sin\theta$$

All the perpendicular components will cancel each other as shown in the Fig. 3.63. Only the parallel components will add up. It is noted that the component of field intensity in the field direction due to charges on differential surface area *ds* is given by Eqn. (3.37). Therefore,

/

The component in the field direction due to charges on dA is

$$dE_4 = \left(\frac{P\cos\theta}{4\pi\varepsilon_0 r^2}\right)\cos\theta \, dA = \left(\frac{P\cos^2\theta}{4\pi\varepsilon_0 r^2}\right) dA$$
$$= \left(\frac{P\cos^2\theta \, ds}{4\pi\varepsilon_0 r^2}\right)(2\pi r^2\sin\theta \, d\theta)$$
$$= \frac{P}{2\varepsilon_0}(\cos^2\theta\sin\theta \, d\theta)$$

Thus, the magnitude of field intensity at A due to charges on the whole spherical surface is

$$E_4 = \int_0^{\pi} \frac{P}{2\varepsilon_0} (\cos^2 \theta \sin \theta \, d\theta)$$
$$= \frac{P}{2\varepsilon_0} \times \frac{2}{3} = \frac{P}{3\varepsilon_0} \qquad \left(\operatorname{since} \int_0^{\pi} \cos^2 \theta \sin \theta \, d\theta = \frac{2}{3} \right)$$

Hence, $E_4 = \frac{P}{3\varepsilon_0}$

In vector form, $\overline{E}_4 = \frac{\overline{P}}{3\epsilon_0}$

We know that

$$\overline{E}_{\rm int} = \overline{E}_1 + \overline{E}_2 + \overline{E}_3 + \overline{E}_4 = \overline{E} + \frac{\overline{P}}{\varepsilon_0} + \frac{-\overline{P}}{\varepsilon_0} + 0 + \frac{\overline{P}}{3\varepsilon_0}$$

i.e.,

 $\overline{E}_{\rm int} = \overline{E} + \frac{\overline{P}}{3\varepsilon_0}$

The above equation is the Lorentz relation for local field. It is seen that $\overline{E}_{int} > \overline{E}$.

EXAMPLE 3.19

A solid contains 5×10^{28} atoms/m³ each with a polarizability of 2×10^{-40} F·m². Assuming that the internal field is given by Lorentz formula. Calculate the ratio of internal field to the external field. $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m).

Solution Given $\alpha_e = 2 \times 10^{-40} \text{ F} \cdot \text{m}^2$, $N = 5 \times 10^{28} \text{ atoms/m}^3$ and $\varepsilon_0 = 8.854 \times 10^{-12} \text{ F/m}$.

Using Lorentz relation for internal field,

$$E_{i} = E + \frac{P}{3\varepsilon_{0}} = E + \frac{N\alpha_{e}E_{i}}{3\varepsilon_{0}}$$
$$E_{i} \left[1 - \frac{N\alpha_{e}}{3\varepsilon_{0}} \right] = E$$

i.e.,

Hence, the ratio of internal field to the external field is

$$\frac{E_i}{E} = \frac{1}{\left(1 - \frac{N\alpha}{3\varepsilon_0}\right)} = \frac{1}{\left(1 - \frac{5 \times 10^{28} \times 2 \times 10^{-40}}{3 \times 8.854 \times 10^{-12}}\right)} = 1.613$$

3.19 CLAUSIUS-MOSOTTI RELATION

Considering the cubic structure, there are no ions or permanent dipoles.

Hence, $P_i = 0$, $P_0 = 0$ and only P_e exists.

Therefore,

 $\overline{P} = N\alpha_e \overline{E}_{int} \qquad (\text{since } \overline{P} = Np_e = N\alpha_e \overline{E})$ $\overline{P} = N\alpha_e \left[\overline{E} + \frac{\overline{P}}{3\varepsilon_0}\right] \qquad \left(\text{since } \overline{E}_{int} = \overline{E} + \frac{\overline{P}}{3\varepsilon_0}\right)$ $\overline{P} \left[1 - \frac{N\alpha_e}{3\varepsilon_0}\right] = N\alpha_e \overline{E}$

i.e.,

 $\overline{P} = \overline{D} - \varepsilon_0 \overline{E}$

Hence,

 $\overline{P} = \frac{N\alpha_e \overline{E}}{1 - \frac{N\alpha_e}{3\varepsilon_0}}$

 $\overline{P} = \varepsilon_0 \varepsilon_r \overline{E} - \varepsilon_0 \overline{E} \qquad (\text{since } \overline{D} = \varepsilon \overline{E} = \varepsilon_0 \varepsilon_r \overline{E})$

We know that, $\overline{D} = \overline{P} + \varepsilon_0 \overline{E}$

i.e.,

or

i.e., $\frac{\overline{P}}{\overline{E}} = \varepsilon_0 \varepsilon_r - \varepsilon_0 = \varepsilon_0 (\varepsilon_r - 1)$

Hence, $\overline{P} = \varepsilon_0 \overline{E} (\varepsilon_r - 1)$

From Eqns. (3.38) and (3.39), we have

$$\frac{N\alpha_{e}\overline{E}}{1-\frac{N\alpha_{e}}{3\varepsilon_{0}}} = \varepsilon_{0}\overline{E}(\varepsilon_{r}-1)$$

$$1-\frac{N\alpha_{e}}{3\varepsilon_{0}} = \frac{N\alpha_{e}}{\varepsilon_{0}(\varepsilon_{r}-1)}$$

$$\frac{N\alpha_{e}}{3\varepsilon_{0}} + \frac{N\alpha_{e}}{\varepsilon_{0}(\varepsilon_{r}-1)} = 1$$

$$\frac{N\alpha_{e}}{3\varepsilon_{0}} \left[1+\frac{3}{\varepsilon_{r}-1}\right] = 1$$

$$\frac{N\alpha_{e}}{3\varepsilon_{0}} = \frac{1}{1+\frac{3}{\varepsilon_{r}-1}} = \frac{\varepsilon_{r}-1}{\varepsilon_{r}+2}$$

Therefore,

$$\left[\frac{\varepsilon_r - 1}{\varepsilon_r + 2}\right] = \frac{N\alpha_e}{3\varepsilon_0}$$
(3.40)

(3.38)

(3.39)

Here, N is the number of moles per volume. Equation (3.40) is known as Clausius-Mosotti equation. From Eqn. (3.40), α_e and ε_r can be determined.

If there are different types of molecules in the dielectric, then

$$\left\lfloor \frac{\varepsilon_r - 1}{\varepsilon_r + 2} \right\rfloor = \frac{1}{3\varepsilon_0} \sum_i N_i \alpha_{ei}$$

where N_i and α_{ei} are appropriate quantities for the type of molecule or atom. Also, from Eqn. (3.40),

$$\left[\frac{\varepsilon_r - 1}{\varepsilon_r + 2}\right] = \frac{N\alpha_e}{3\varepsilon_0}$$

$$\varepsilon_r = \frac{1 + \frac{2}{3\varepsilon_0} N\alpha_e}{1 - \frac{1}{3\varepsilon_0} N\alpha_e}$$

It can be written as

$$\left(\frac{N\alpha_e}{3\varepsilon_0}\right)\left(\frac{M}{\rho}\right) = \left(\frac{\varepsilon_r - 1}{\varepsilon_r + 2}\right)\left(\frac{M}{\rho}\right)$$

Now $\frac{NM}{\rho} = N_A$, the Avogadro number. Therefore,

$$\frac{N_A \alpha_e}{3\varepsilon_0} = \left(\frac{M}{\rho}\right) \left(\frac{\varepsilon_r - 1}{\varepsilon_r + 2}\right) \tag{3.41}$$

Here, ε_r is the macroscopic quantity and α_e is the microscopic quantity. Equation (3.41) relates these quantities and by measuring ε_r , α_e can be found.

In case of optical materials, if *n* is the refractive index of the material, then $\varepsilon_r = n^2$ and the Clausius-Mosotti relation in Eqn. (3.40) becomes

$$\left[\frac{n^2 - 1}{n^2 + 2}\right] = \frac{N\alpha_e}{3\varepsilon_0}$$

EXAMPLE 3.20

The atomic weight and density of sulphur are 32 and 2.08 g/cm³ respectively. The electronic polarizability of the atom is 3.28×10^{-40} F·m². If sulphur solid has cubic symmetry, then determine its relative permittivity.

Solution Given M = 32, $\rho = 2.08 \text{ g/cm}^3 = 2.08 \times 10^3 \text{ kg/m}^3$ and $\alpha_e = 3.28 \times 10^{-40} \text{ F} \cdot \text{m}^2$.

Using Clausius-Mosotti relation,

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{N\alpha_e}{3\varepsilon_0} = \frac{N_A \rho \alpha_e}{3M\varepsilon_0}$$
$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{(6.023 \times 10^{26})(2.08 \times 10^3)(3.28 \times 10^{-40})}{3 \times 32 \times 8.85 \times 10^{-12}} = 0.483$$
$$\varepsilon_r - 1 = 0.483(\varepsilon_r + 2) = 0.483\varepsilon_r + 0.966$$

i.e., or

 $\mathcal{E}_r(1-0.483) = 1+0.966$

Therefore,

$$\varepsilon_r = \frac{1.966}{0.517} = 3.8$$

EXAMPLE 3.21

A dielectric material has $\varepsilon_r = 4.94$ and $n^2 = 2.69$. Calculate the ratio between electronic and ionic polarizability of this material.

Solution Given $\varepsilon_r = 4.94$ and $n^2 = 2.69$.

Using Clausius-Mosotti relation,

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{N\alpha}{3\varepsilon_0} = \frac{N(\alpha_e + \alpha_i)}{3\varepsilon_0} \qquad \text{(since } \alpha_0 \text{ is very small and negligible)}$$

$$\frac{N(\alpha_e + \alpha_i)}{3\varepsilon_0} = \frac{4.94 - 1}{4.94 + 2} = 0.568 \qquad (1)$$

(2)

At optical frequencies, $\varepsilon_r = n^2$

Therefore,

$$\frac{n^2 - 1}{n^2 + 2} = \frac{N\alpha_e}{3\varepsilon_0}$$

or

Dividing Eqn. (1) by Eqn. (2), we get

$$\frac{N(\alpha_e + \alpha_i)}{N\alpha_e} = \frac{0.568}{0.360} = 1.578$$

 $\frac{N\alpha_e}{3\varepsilon_0} = \frac{2.69 - 1}{2.69 + 2} = 0.360$

or

$$\frac{\alpha_i}{\alpha_e} = 0.578$$

Therefore,
$$\frac{\alpha_e}{\alpha_e} = \frac{1}{1}$$

$$\frac{\alpha_e}{\alpha_i} = \frac{1}{0.578} = 1.73$$

 $1 + \frac{\alpha_i}{\alpha} = 1.578$

FREQUENCY DEPENDENCE OF POLARIZATION 3.20

In many practical situations, a dielectric is subjected to an alternating electric field. Hence, in case of nonpolar dielectric, the induced dipoles oscillate in response to the frequency of the applied field and in polar dielectrics, the permanent dipoles try to reorient in response to the applied field. Thus, it is necessary to know about the response of each polarization in different frequency regions.

On application of an electric field, the polarization process occurs as a function of time. The polarization P(t)as a function of time *t* is given by

$$P(t) = P(1 - e^{-t/t_r})$$

where P is the maximum polarization attained on prolonged application of a static field and t_r is the relaxation time for the particular polarization process.

The relaxation time t_r is the time taken for the polarization to reach 0.63 of the maximum value P. This varies widely for different polarization processes.

Electronic polarization is extremely rapid. Even when the frequency of the applied voltage is very high in the optical range (~ 10¹⁵ Hz), electronic polarization occurs during every cycle of the applied voltage.



Fig. 3.65 Frequency dependence of polarization processes

On application of a field, the ions take a longer time to respond and get displaced as they are heavier than the electrons. Hence, ionic polarization is slower than electronic polarization. The response time is of the order of 10^{-13} sec. Hence, there is no ionic polarization at optical frequencies. When the frequency of the applied voltage is less than 10^{13} Hz, the ions respond and both electronic and ionic polarization will occur.

Orientation polarization is slower than ionic polarization. The response time for orientation polarization is around 10^{-10} sec and occurs when the frequency of the applied voltage is in the audio range.

Space charge polarization takes the longest time to set in. The response time is very large and is of the order of 10^{-2} sec. This occurs at power frequencies (50-60 Hz).

Figure 3.65 and Table 3.7 illustrate all the types of polarization in different frequency ranges. At optical frequencies ($\sim 10^{15}$ Hz), only electronic polarization occurs. At $\sim 10^{13}$ Hz, ionic polarization and electronic polarization occur. In the range 10^6 to 10^{10} Hz, orientation polarization gets added and at 10^2 Hz range, space charge polarization also contributes.

Freq	uency range in Hz	Type of polarization
~1013	5	electronic
~1012	3	ionic and electronic
~10 ⁶	-10^{10}	orientation, ionic and electronic
~10 ²		space-charge, orientation, ionic and electronic

 Table 3.7
 Types of polarization in different frequency ranges

3.21 EFFECT OF TEMPERATURE ON DIELECTRIC CONSTANT

Among all the types of polarization, only orientational and space change polarization depends on temperature.

In orientational polarization, thermal energy decreases the tendency of the permanent dipoles to align themselves along the applied field. This results in a decrease in the dielectric constant with increasing temperature.

In space charge polarization, the increase of temperature facilitates the diffusion of ions. Thermal energy may also aid in overcoming the activation barrier for the orientation of polar molecules in the direction of the field. Hence, the relative dielectric constant may increase with increase of temperature. Thus, the overall dielectric constant will be the net effect of temperature on orientational and space charge polarization.

3.22 DIELECTRIC LOSS

Electrical energy is absorbed by a dielectric material and is dissipated in the form of heat when it is placed in an alternating field. This dissipation of energy is called dielectric loss. Dielectric loss can occur in both direct and alternating voltages. It is less in direct voltage than that in alternating voltage. Since this involves heat generation and heat dissipation, it finds applications in food processing and in high voltage applications.

When an AC voltage is applied to a perfect dielectric like vacuum or purified gas, it does not absorb electrical energy and hence, there is no loss of electrical energy as shown in Fig. 3.66(a). Polarization of the dielectric is in phase with the voltage. In such a case, the charging current leads the applied voltage by an angle of 90° as shown in Fig. 3.66(a). This happens when the period of the applied voltage is much larger than the relaxation time of a polarization process and polarization is essentially complete at any instant during each cycle.

We know that power loss,

$$P_L = VI \cos \theta$$

When $\theta = 90^\circ$, $P_L = VI \cos 90^\circ = 0$

This means that there is no power loss in an ideal or perfect dielectric.

However, a practical dielectric has always some loss of the electrical energy in the form of heat. In this case, the leakage current does not lead the applied voltage exactly by 90° .

When the period is in the same range as the relaxation time, resonance occurs. Here, the current leads the voltage by $(90^{\circ} - \delta)$ as illustrated in Fig. 3.66 (b). Here, δ is called the loss angle and tan δ is taken as a measure of the electrical loss due to resonance and is known as loss tangent. The current can be factorized into a component at 90° to the voltage, as in an ideal capacitor, and another component parallel to the voltage. The parallel component is the real part and results in $I^2 R$ loss.

Loss tangent can be also expressed as $\tan \delta = \frac{\varepsilon_r''}{\varepsilon_r'}$

where ε' and ε''_r are real and imaginary parts of relative permittivity as shown in Fig. 3.66(c).



Fig. 3.66 Energy loss in dielectric (a) No energy loss when period of ac voltage does not match relaxation time (b) when there is a match, energy loss occurs with the current leading the voltage by less than 90° and (c) current density components with respect to the field direction

Expression for Dielectric Loss (or Loss Tangent)

The power loss in a dielectric having a capacitance C for applied voltage V of frequency f is given by

 $P_L = VI \cos \theta$

Since $\theta = 90^{\circ} - \delta$, we have

$$P_{L} = VI \cos (90^{\circ} - \delta)$$

$$P_{L} = VI \cos \delta \qquad [\text{since } \cos (90^{\circ} - \delta) = \sin \delta] \qquad (3.42)$$

Since the reactance of capacitor is X_C , we have

$$I = \frac{V}{X_C} \tag{3.43}$$

Further, X_C depends on frequency of applied AC voltage and capacitance. It is given by

$$X_C = \frac{1}{2\pi fC} \tag{3.44}$$

Substituting Eqn. (3.44) in (3.43), we have

$$I = \frac{V}{1/2\pi fC} = V2\pi fC$$
(3.45)

Substituting Eqn. (3.45) in (3.42), we have

$$P_L = V(V2\pi f)C\sin\delta = 2\pi fCV^2\sin\delta$$

In most of the dielectrics, the angle δ is very small i.e., $\sin \delta = \tan \delta$.

Therefore, the dielectric power loss is

$$P_L = 2\pi f C V^2 \tan \delta$$

where $\tan \delta$ is called the power factor of the dielectric.



Fig. 3.67 Variation of power loss with frequency

The power loss is dependent on tan δ as long as the other factors like voltage, frequency and capacitance are constants. Otherwise, power loss changes with frequency and applied voltage. Its value is high at electrical frequencies and low at optical frequencies as shown in Fig. 3.67. The dielectric loss is also affected by rise in temperature and humidity.

The dielectric losses in the radio frequency region are usually due to dipole rotation or due to ions jumping from one equilibrium position to another. The dielectric losses in the optical region, known as optical absorption, are due to electronic polarization. The dielectric losses at lower frequencies are mainly due to DC resistivity. The dielectric losses associated with the ionic vibrations, the frequencies of which fall in the infrared region, are usually referred to as infrared absorption.

3.23 DIELECTRIC BREAKDOWN

When the strength of electric field applied to a dielectric exceeds a critical value, a very large current flows through it. The dielectric looses its insulating property and becomes a conductor. This phenomenon is known as *dielectric breakdown*.

This is because, a considerable number of covalent bonds break and electrons may get excited to energies within the conduction band. These electrons acquire a large kinetic energy and cause localized melting, burning and vapourization of material leading to irreversible degradation and failure of the material leading to breakdown.

3.23.1 Dielectric Strength

The dielectric strength of a dielectric is a measure of the maximum electric field that the material can withstand under ideal conditions without breakdown.

Dielectric strength,
$$E_{\text{max}} = \frac{\text{Dielectric breakdown voltage}}{\text{Thickness of the dielectric}} = \frac{V_{\text{max}}}{d}$$

The dielectric strength depends on the thickness of the insulating material and the amount of time for which the dielectric is subjected to the electric field.

It is the limiting field intensity above which a breakdown occurs and hence, it will lose its charge storage property. The dielectric strength will get decreased due to various parameters such as moisture, contamination, elevated temperature, ageing and mechanical stress.

3.23.2 Breakdown Mechanisms in Solid Dielectrics

The physical pattern of breakdown of solid dielectrics may differ in various cases. However, the fundamental breakdown mechanisms are as follows:

- 1. Intrinsic and Avalanche breakdown
- 2. Thermal breakdown
- 3. Discharge breakdown
- 4. Electrochemical breakdown
- 5. Defect breakdown

1. (a) Intrinsic Breakdown

Beyond a critical value of electrical field strength, the charge displacement increases in a dielectric and there is an electrical breakdown due to physical deterioration of the material. When the applied electrical field is very large, some electrons in the valence band move to the conduction band due to breaking of covalent bonds. As a result, a large current flows through the dielectric and breakdown occurs. This type of breakdown is called intrinsic breakdown.

(b) Avalanche Breakdown

The conduction electrons are accelerated with very high velocities and energy on further increase in the electrical field. They collide with valence electrons in the covalent bond thereby, breaking them and releasing more and more electrons into the conduction band. An avalanche of electrons is formed in the process resulting in enormous current flow and breakdown occurs. This type of breakdown is called avalanche breakdown. Intrinsic and avalanche breakdowns require large electrical fields and occur even at low temperatures. They can occur in thin samples and within a short span of time i.e. microseconds.

2. Thermal breakdown

In dielectric materials, energy due to the dielectric loss appears as heat. This heat must be dissipated away to the surroundings. If the rate of heat generation is larger than the rate of heat dissipation, the temperature of the dielectric increases which results in local melting. Eventually the dielectric property breaks down and this type of breakdown is called thermal breakdown.

The characteristics of thermal breakdown are as follows:

- (i) It occurs at high temperatures.
- (ii) The breakdown time is of the order of milliseconds.
- (iii) In AC fields, the breakdown strength lowers.
- (iv) The breakdown strength depends on the size and shape of the material.
- (v) It requires moderate electrical fields.

3. Discharge Breakdown

Breakdown of dielectrics by gas discharge is classified as external breakdown or internal breakdown.

External breakdown is caused by a glow or corona discharge and is observed at sharp edges of electrodes. The discharge causes gradual deterioration of the solid dielectric held between the electrodes. Such deterioration is accompanied by the formation of carbon. Therefore, the damaged areas become conducting. Eventually, conducting paths are formed leading to a powerful arc and total failure of the dielectric. Such breakdown is caused mainly due to the contamination of the dielectric surface by conducting impurities such as dust, moisture etc.,

Internal breakdown occurs when a dielectric contains occluded gas bubbles. When this type of dielectric is subjected to electrical field, the gases present in the material will easily ionize and thus, produce large ionization current. This causes electrical deterioration and leads to dielectric breakdown.

The characteristics of thermal breakdown are as follows:

- (i) This discharge breakdown occurs at low voltages.
- (ii) It occurs in the dielectric material where there are a large number of occluded gas bubbles.
- (iii) When discharge takes place at a point, the surrounding places are burnt and hence, their electrical properties are affected.
- (iv) It depends on the frequency of applied voltage.

4. Electrochemical Breakdown

Electrochemical breakdown is analogous to thermal breakdown. Many materials have free ions which cause leakage current in the presence of electric field. When temperature increases, mobility of ions increases and also increases leakage current. Hence, electrochemical reaction takes place in the material. The field induces chemical reactions which reduce the resistance of the dielectric. This leads to leakage current and energy loss in the dielectric material and finally dielectric breakdown occurs.

The characteristics of electrochemical breakdown are as follows:

- (i) It depends on the concentration of ions and magnitude of leakage current.
- (ii) It occurs at ordinary temperatures.

5. Defect Breakdown

If the surface of the dielectric material has defects such as cracks and porosity, impurities such as dust or moisture may deposit at these defects. These impurities lead to breakdown.

Remedies for dielectric breakdown To avoid breakdown, the insulating material should have the following properties:

- (i) High resistivity to reduce leakage current
- (ii) Small dielectric loss
- (iii) Sufficient mechanical strength
- (iv) Resistance to oils, liquids, gas fumes, acids and alkalies
- (v) Small thermal expansion to prevent mechanical damage
- (vi) Fire-proof properties

3.24 HIGH-K DIELECTRICS-FERROELECTRICITY

Materials which exhibit electric polarization even in the absence of the applied electric field are known as Ferroelectric materials. Ferroelectricity is a result of dielectric hysteresis. Since these materials exhibit hysteresis effects similar to those observed in ferromagnetic materials, they are called ferroelectric materials.

These materials have permanent dipole moment in each atom or molecule. The dielectric constants of these materials are several orders of magnitude larger than that in ordinary dielectrics. Ferroelectricity refers to the creation of enormous value of induced dipole moment in a weak electric field as well as existence of electric polarization even in the absence of applied electric field. Hence, these are sometimes referred to as high-k dielectrics.

Ferroelectricity was first discovered in Rochelle salt (Na KC_4 H₄ O₆.4H₂O). It exhibits spontaneous polarization over a range of temperature -18° C to 22° C.

Barium titanate (BaTiO₃), Potassium dihydrogen phosphate (KH₂PO₄), Ammonium dihydrogen phosphate (NH₄ H₂ PO₄), and Lithium Niobate (Li NbO₃. 4H₂O) are some other typical examples of ferroelectric materials.

Properties of Ferroelectric Materials

- (i) Ferroelectric materials can easily be polarized even by very weak electric fields.
- (ii) They possess very high values of dielectric constant or relative permittivity which is of the order of 1000 to 10,000.
- (iii) The static dielectric constant of ferroelectric materials change with temperature according to the relation given by

$$\varepsilon = \frac{C}{T - T_C} \quad (T > T_C) \tag{3.46}$$

The above equation is known as Curie-Weiss law. Here, C is called the Curie constant and T_C is the Curie temperature. The variation of ε_r with temperature in barium titanate is shown Fig. 3.68.



Fig. 3.68 Variation of dielectric constant with temperature in barium titanate

(iv) They possess spontaneous electric polarization, i.e., polarization without the help of an external electric field. However, the spontaneous polarization occurs only within a definite temperature range and up to T_C i.e., for $T \le T_C$, where T_C is the ferroelectric Curie temperature. When $T > T_C$, they are converted into paraelectric materials.

Condition for Spontaneous Polarization

We know that
$$\overline{P} = \frac{N\alpha_e \overline{E}}{1 - \frac{N\alpha_e}{3\varepsilon_0}}$$

If the denominator becomes zero, a polarization catastrophe is obtained. This can be interpreted as that there is polarization even when there is no external field or there is spontaneous polarization just as there is spontaneous magnetization in ferromagnetics. The condition for spontaneous polarization is given by

$$\frac{N\alpha_e}{3\varepsilon_0} = 1$$

Hence, with $\frac{N\alpha_e}{3\varepsilon_0} = 1$, the dielectric constant will become infinite, but it is concluded that the materials

will become spontaneously polarized. The variation of spontaneous polarization with temperature in barium titanate is shown in Fig. 3.69.

(v) In a ferroelectric, the dielectric polarization depends nonlinearly on the applied electric field. In ordinary dielectrics, the polarization varies linearly with the applied electric field. Because of this, ferroelectrics are known as nonlinear dielectrics. They exhibit hysteresis under the action of an alternating voltage. Lagging of polarization behind the applied electric field is called dielectric hysteresis. The polarization versus electric field curve is known as a *ferroelectric hysteresis loop* as shown in Fig. 3.70.



Fig. 3.69 Spontaneous polarization in barium titanate



С

E

- (vi) They exhibit domain structure as in the case of ferromagnetic materials.
- (vii) Ferroelectric materials exhibit Piezoelectricity and Pyroelectricity. Piezoelectricity means the creation of electric polarization by mechanical stress. Pyroelectricity means the creation of electric polarization by thermal stress.

Piezoelectric materials are used to make pressure transducers, ultrasonic transducers and microphones. Examples: (a) Quartz, Lithium Niobate and Barium Titanate among the crystalline materials (b) Lead Zirconium Titanate, Calcium Barium Titanate and Lead Barium Niobate among the ceramic materials *Pyroelectric materials* are used to make high sensitive infrared detectors. Examples: Barium Titanate, Triglycine sulphate, Lithium Niobate, Lithium Tantalite and Polyvinyl fluoride.

- (vii) Some ferroelectric semiconductors like Ba Ti O_3 Sr Ti O_3 , Ba Ti O_3 Pb Ti O_3 and Sr Ti O_3 Pb Ti O_3 are used to make posistors which are used to measure and control temperature like thermistors. These have positive temperature coefficient of resistance.
- (viii) Electrets are ferroelectric materials and are electrostatic analog of permanent magnets. Electrets possess a gross permanent electric dipole moment. They are manufactured from certain types of waxes, plastics and ceramics. When these waxes are subjected to high electric fields ($\approx 10^6$ V/m), they are polarized in their molten state and retain a permanent polarization after solidifying, even though the external polarizing field is removed. Electrets are used in capacitor microphones and gas filters to capture submicron particles by electrostatic attraction. Further, the electret bondages are used over the fractured bones to speed up the healing process.

Spontaneous polarization and domains in BaTiO₃

Figure 3.71 shows the 3-D view and cross-sectional view of the cubic crystal structure of $BaTiO_3$ above 120°C. The structural changes in $BaTiO_3$ crystal due to lattice variation give rise to ferroelectricity. Above 120°C, $BaTiO_3$ has a cubic crystal structure with the titanium ions exactly at the body centres, Barium ions are at the corners and Oxygen ions are at the face centres. This is the unpolarized phase of the ferroelectric solid, and as the polarization vectors of different domains orient in different directions



Fig. 3.71 Cubic crystal structure of BaTiO₃ above 120°C (a) 3-D view and (b) cross-sectional view

and the net polarization of the solid is zero. Hence, there is no spontaneous dipole moment as shown in Fig. 3.71.

If the crystal is cooled below 120°C, the titanium ions shift to one side of the body centre. This displaces the neighbouring ions along one of the crystal directions which become elongated and the structure becomes tetragonal. The energy of the crystal is lower when the Ti⁴⁺ ion in each unit cell is slightly displaced along the corner direction, as shown in Fig. 3.72, which generates a dipole moment in each unit cell. Strong interactions between the adjacent permanent dipoles lower the energy of the whole crystal when they are all aligned in the same direction within some volume of the solid. Such regions are said to be spontaneously polarized and are known as ferroelectric domains. Here, the distortion of the crystal that takes place when spontaneous polarization occurs just below T_C is very small relative to the dimensions of the unit cell. For BaTiO₃, for example, c/a is 1.01 and the displacement of the Ti⁴⁺ ion from the center is only 0.012 nm, compared with a = 0.4 nm.

Similarly, at 5°C spontaneous polarization direction corresponds to the face diagonal direction and at -80°C, the direction corresponds to a body diagonal. At these temperatures, there is an enormous value for dielectric constant. In fact, BaTiO₃ has more transition temperatures.



Fig. 3.72 Structure of BaTiO₃ (a) below its Curie temperature 120°C which is slightly tetragonal (b) Formation of electric dipole moment in unit cell due to a slight shift of the central ion Ti⁴⁺ with respect to the surrounding O²⁺ ions of the unit cell

Applications of Ferroelectric Materials

(i) Ferroelectric Energy Converter

Ferroelectric crystals exhibit the pyroelectric effect. i.e., a change in the temperature of the crystal produces a change in its polarization. Using this effect, heat energy can be converted into electrical energy. The efficiency of this conversion is about 5.6 %. First, the capacitor with Ba TiO₃ is charged to a voltage of V_1 at a temperature T_1 just below the Curie temperature θ_f as shown in Fig. 3.73(a).

Now $q = C_1 V_1$ where C_1 is the capacitance of the capacitor.

Initial energy $=\frac{1}{2}qV_1$ = area of the triangle *abc* as shown in Fig. 3.73(b).

The capacitor is isolated from the battery and it is heated up o T_2 . Therefore, the dielectric constant ε_r decreases for Ba TiO₃ when the temperature is increased from 15°C to 30°C,



Fig. 3.73 Ferroelectric energy conversion

 $\frac{\text{Value of dielectric constant at 30^{\circ} C}}{\text{Value of dielectric constant at 15^{\circ} C}} = \frac{\varepsilon_{r1}}{\varepsilon_{r2}} = 5$

Since the dielectric constant decreases, the capacitance also decreases and at temperature T_2 , the capacitance of the capacitor is equal to C_2 .

i.e., $C_2 < C_1$

Since the charge q in the capacitor is constant and $C_2 < C_1$,

$$q = C_1 V_1 = C_2 V_2$$
$$V_2 > V_1$$
$$V_2 = \frac{V_1 C_1}{C_2} = \frac{V_1 \varepsilon_{r1}}{\varepsilon_{r2}} = 5V_1$$

or

Hence, the increase in energy, $\Delta W = \frac{1}{2}q(V_2 - V_1)$

The capacitor is discharged through a load resistor R and a rechargeable battery of voltage V_1 to get electrical power. The capacitor should be cooled to T_1 , to complete the cycle.

Applications

- The high dielectric constant of ferroelectric crystals is also useful for storing energy in small sized capacitors in electrical circuits.
- (ii) In optical communication, the ferroelectric crystals are used for optical modulation.
- (iii) These are used in electro acoustic transducers such as microphone.
- (iv) Ferroelectric crystals exhibit the piezoelectric property which finds applications in transducers.

TWO MARKS QUESTIONS AND ANSWERS

1. Define Bohr magneton.

The orbital magnetic moment and the spin magnetic moment of an electron in an atom can be expressed in terms of atomic unit of magnetic moment which is called Bohr magneton. It is represented by

$$\mu_B = \frac{e\hbar}{2m_e} = 9.27 \times 10^{-24} \,\mathrm{A \cdot m^2}$$

2. What is meant by magnetic moment?

The magnetic dipole moment (\overline{m}) of the loop is defined as the product of current *I* and area *A* of the loop with its direction normal to the loop. Its magnitude is given by m = IA.

3. Define magnetization (or) intensity of magnetization.

Magnetization is the process of converting a non-magnetic material into a magnetic material. It measures the magnetization of the magnetized specimen. Also, the magnetic moment (*m*) per unit volume (*V*) of a solid is called magnetization. It is given by M = m/V.

4. Define magnetic field.

The space around the magnet (or) the current carrying conductor where the magnetic effect is felt is called magnetic field. Magnetic field is produced by permanent magnets like horse shoe magnets and temporarily by electromagnets or superconductor magnets. Since the magnetic field is a vector quantity, it is denoted by \overline{H} . The unit of \overline{H} is ampere-turns per metre (A/m) in SI system.

5. What is meant by flux density? How is it related to magnetic field intensity?

The number of field or flux lines (ϕ) passing normally through unit area of cross-section is called the magnetic flux density \overline{B} . The unit of flux density is Wb/m² or Tesla. It is given by

$$\overline{B} = \frac{\text{Magnetic flux}}{\text{area}} = \frac{\phi}{A}$$

It is related to field intensity by $\overline{B} = \mu \overline{H}$ where μ is the permeability of the medium.

6. What is magnetic susceptibility?

The magnetic susceptibility χ_m is defined as the magnetization \overline{M} produced in the material per applied magnetic field \overline{H} . It is given by $\chi_m = \frac{\overline{M}}{\overline{H}}$

7. Define magnetic permeability.

Magnetic permeability is defined as the ratio of the magnetic flux density \overline{B} to the magnetic field intensity \overline{H} . i.e., $\mu = \mu_0 \mu_r = \overline{B}/\overline{H}$

8. Define relative permeability.

Relative permeability is defined as the ratio of the permeability of the medium μ to the permeability of free space μ_0 . It is given by $\mu_r = \mu/\mu_0$

9. What are different sources of permanent magnetic moment?

The different sources of permanent magnetic moment are

- (i) Orbital magnetic moment of the electrons
- (ii) Spin magnetic moment of the electrons, and

3.70 Physics for Electronics Engineering and Information Science

(iii) Spin magnetic moment of the nucleus

10. Classify different types of magnetic materials based on magnetic moment.

Based on magnetic moment, magnetic materials are classified into four different types. They are:

- (i) Those not having any permanent magnetic moment are known as diamagnetic materials.
- (ii) If permanent dipoles do not interact among themselves, the materials are known as paramagnetic materials.
- (iii) If the permanent dipoles are strong such that all the dipoles line up in parallel, the materials are known as ferromagnetic materials.
- (iv) If the permanent dipoles line up in anti-parallel direction, the materials are known as antiferromagnetic or ferrimagnetic materials.

11. What is diamagnetism?

When an external magnetic field is applied, the electrons align perpendicular to the field direction and their magnetic moments oppose the external magnetic field. Since the motion of the orbital electrons changes, the induced magnetic moment is in a direction opposite to the applied field direction. Due to this effect, the material gets very weakly repelled in the magnetic field.

12. Define ferromagnetism.

The materials which have finite values of magnetization even in the absence of applied magnetic field are called ferromagnetic materials. This phenomenon is called ferromagnetism. Examples: Fe, Co & Ni. All these materials exhibit high degree of magnetization.

13. What is anti-ferromagnetism?

In anti-ferromagnetism, electron spin of neighboring atoms are aligned in anti-parallel. The susceptibility is small and positive in anti-ferromagnetic materials and it depends greatly on temperature.

14. What is a soft magnetic material?

Material that can easily be magnetized and demagnetized is called soft magnetic material. Examples: Iron-silicon alloys and Iron-cobalt alloys.

15. What is a hard magnetic material?

Material that is very difficult to magnetize and demagnetize is called hard magnetic material. Examples: Tungsten steel and Carbon steel.

16. Differentiate hard and soft magnetic materials.

S. No.	Hard magnetic materials	Soft magnetic materials
1.	Difficult to magnetize and demagnetize	Easy to magnetize and demagnetize
2.	Low Permeability	High Permeability
3.	High coercivity	Low coercivity
4.	High remanence	Low remanence
5.	High hysteresis loss	Low hysteresis loss

17. List the properties of diamagnetic materials.

The properties of diamagnetic materials are:

- (i) Permanent dipoles are absent
- (ii) Effect is weak and often masked by other kind of magnetism
- (iii) When placed inside a magnetic field, magnetic lines of force are repelled
- (iv) Magnetic susceptibility is negative
- (v) Magnetic susceptibility is independent of applied magnetic field strength
- (vi) Relative permeability is slightly less than unity

18. List the properties of paramagnetic materials.

The properties of paramagnetic materials are:

- (i) They possess permanent magnetic dipoles.
- (ii) These dipoles are non-interacting.
- (iii) The dipoles are randomly oriented and hence, in the absence of external applied magnetic field, the net magnetization in any given direction is zero.
- (iv) When placed inside a magnetic field, it attracts the magnetic lines of force.
- (v) Paramagnetic susceptibility is positive and depends greatly on temperature.

19. Mention the differences between diamagnetic and paramagnetic materials.

The differences between diamagnetic and paramagnetic materials are given below.

<i>S. No.</i>	Diamagnetism	Paramagnetism
1.	Permanent dipoles are absent.	Permanent dipoles are present.
2.	When placed in the magnetic field, the magnetic lines of forces are repelled.	Magnetic lines of forces are attracted.
3.	Magnetic susceptibility is negative	Magnetic susceptibility is positive

20. Distinguish between ferro and antiferro magnetic materials.

The comparison of ferro and antiferro magnetic materials is given below.

S. No.	Ferromagnetism	Antiferromagnetism		
1.	All spin magnetic moments are aligned parallel in the same direction.	Spin magnetic moments are aligned antiparallel.		
2.	Susceptibility is positive and very large.	Susceptibility is positive and small.		

21. How will you differentiate magnetic materials based on their spin alignment?

The comparison of magnetic materials based on their spin alignment is given below.

No spin	Random spin alignment	Parallel spin alignment	Anti-parallel spin alignment
Diamagnetic materials	Paramagnetic materials	Ferromagnetic materials	Anti-ferromagnetic materials

22. On the basis of spin, how the materials are classified as para, ferro, antiferro and ferrimagnetic.

- (i) Paramagnetic materials have few unpaired electron spins of equal magnitudes.
- (ii) Ferro magnetic materials have many unpaired electron spins with equal magnitudes.
- (iii) Anti ferro magnetic materials have equal magnitude of spins but in antiparallel manner.
- (iv) Ferrimagnetic materials have spins in antiparallel manner but with unequal magnitudes.

23. What is Curie constant? What is Curie Law?

Langevin showed that classical paramagnetic susceptibility (χ_m) due to the alignment of magnetic moments along field direction is given by

$$\chi_m = \frac{\mu_0 N \mu^2}{3kT} = \frac{C}{T}$$

where C is known as Curie constant and this relation is known as Curie law.

24. Give Curie-Weiss Law and its importance.

According to Curie-Weiss law,

$$\chi_m = \frac{C}{T - \theta}$$

where *C* is Curie constant and θ is the Curie temperature. This law explains the relationship between para and ferromagnetism. When $T < \theta_f$, the material is ferromagnetic and for $T > \theta_f$, it is paramagnetic. (Here, θ_f is the ferromagnetic Curie temperature).

25. What is Heisenberg theory of ferromagnetism?

The molecular field based on simple dipole-dipole interaction was found to be less and hence cannot be accounted for the existence of ferromagnetism. Heisenberg removed this discrepancy by assuming the quantum exchange interaction between the electron spins instead of dipole-dipole interaction.

26. What is meant by domain?

According to Weiss, ferromagnetic materials consist of a number of small regions called domains. The size of a domain varies from 10^{-6} m to entire size of the crystal. A magnetic domain is completely magnetized in single direction. A domain acts as single magnetic dipole and has definite boundaries.

27. What is domain theory of ferromagnetism?

The domain theory of ferromagnetism is the explanation of the structure and hysteresis property of ferromagnetic materials based on the concept of domains proposed by Weiss. Ferromagnetic material consists of a number of regions called domains which are spontaneously magnetized due to parallel alignment of all magnetic dipoles. The direction of spontaneous magnetization varies from domain to domain.

28. Give the two possible ways of domain magnetization.

The two possible ways of domain magnetization are by (i) motion of domain wall and (ii) rotation of domain.

29. Define exchange energy.

The exchange energy is defined as the energy required aligning the atomic magnets into single domain. This work done is stored as potential energy. This energy arises from the interaction of electron spins. This interaction energy makes the adjacent dipoles align themselves.

30. Define crystal anisotropic energy.

The excess energy required per unit volume of the substance to magnetize it along the particular direction over that required to magnetize it along the easy direction is called the crystal anisotropic energy.

31. What is magnetostriction energy?

The change in dimension of ferromagnetic material when it is magnetized is known as magnetostriction. The crystal lattice strain energy associated with magnetostriction is known as magnetostriction energy.

32. Define magnetic domains and domain walls.

According to Weiss, a virgin specimen of ferromagnetic material consists of a number of regions separated by boundary (= 10^6 m or larger) which are spontaneously magnetized. These regions are called domains. The boundaries between different domains are called domain walls.

33. Mention the energies involved in domains of ferromagnetic materials

There are five types of energies involved in domains of ferromagnetic materials. They are

- (i) Exchange energy
- (ii) Magnetostatic energy
- (iii) Anisotropy energy
- (iv) Domain Wall energy or Bloch wall energy
- (v) Magnetostrictive energy

34. Explain the motion of domain walls and the rotation of domains

When a weak external magnetic field is applied to the specimen of a ferromagnetic material, the magnetic moment increases and the dipoles are oriented unfavorably. Thus, the volume of the domain changes and the boundaries of the domain size increases.

When a strong magnetic field is applied to the specimen of a ferromagnetic material, the magnetic moment tends to change their direction towards the applied field direction. Then, due to this alignment, rotation of domains occurs.

35. What is the Magnetic principle of computer data storage?

The soft and hard magnetic properties of materials are made use of in storage of data in computers. The storage of data generally involves the following steps:

- (i) Writing data on a disc or tape (recording medium) which may be in the form of electrical signals.
- (ii) Reading back the data whenever required.

It is known that certain ferromagnetic materials are hard and can be permanently magnetized. Such materials can be used for storing data. The writing/reading heads should be made of a soft magnetic

3.74 Physics for Electronics Engineering and Information Science

material like Permalloy, i.e., materials that can be easily magnetized and demagnetized. They essentially consist of an electromagnet, i.e., a coil wound around a soft magnetic material.

36. What is the principle of Giant Magnetoresistance and a GMR sensor?

Giant Magnetoresistance (GMR) is a result of spin-dependent scattering of charges, which reduces the resistance. It is known from domain theory that the spins of electrons in a ferromagnetic material can be either spin-up or spin-down, based on the component of magnetic field along a chosen axis. Now, when an electron moves into such a magnetized region, it will be scattered depending on its spin. Scattering will be lesser for electrons with spins in the direction of the magnetic field as compared to those with spins opposite to the direction of the magnetic field. This principle is used in a GMR sensor.

37. What are dielectric materials?

Dielectric materials are the insulating materials used to store the electrical energy. It has high resistivity in the order of 10^{10} to $10^{20} \Omega \cdot m$.

38. What are polar and non-polar dielectrics?

If the centre of the positive and negative charges coincides, there will be an intrinsic dipole moment in dielectrics. Such dielectrics are polar dielectrics. Examples: H_2O , HCl.

If the centre of positive and negative charges does not coincide, the net magnetic dipole moment is zero. Such dielectrics are non-polar dielectrics. Examples: H_2 , O_2 , N_2 .

39. How does the dielectric polarization in polar and non-polar substances vary with temperature?

In polar dielectrics, the polarization depends on temperature. It decreases when the temperature increases. The non-polar dielectrics are independent of temperature.

40. What is meant by polarization?

The process of producing induced electric dipoles in dielectrics by an external electric field is called polarization.

41. Mention four types of polarization mechanism that can take place in the presence of an electric field in dielectric materials.

The four types of polarization mechanism that can take place in the presence of an electric field in dielectric materials are (i) Electronic Polarization (ii) Ionic Polarization (iii) Orientation Polarization and (iv) Space charge Polarization

42. Define electronic polarization

The process of displacing central positive nucleus and the negative electron cloud resulting in a dipole moment due to an electric field is called electronic polarization.

43. What is ionic polarization?

Ionic polarization is defined as the process of inducing dipoles by slightly displacing the anions and cations from their rest position by an electric field.

44. Define orientation polarization.

The process of orienting the permanent dipoles of polar molecules along the field direction is known as orientation polarization.

45. What is meant by space charge polarization?

When an electric field is applied, the positive and negative charges will accumulate at the respective electrodes. This process is referred to as space charge polarization.

46. Give the frequency dependence of polarization mechanisms.

The mechanisms of polarization are dependent on frequency and they are as follows:

- (i) At optical frequency (10^{15} Hz) , the electronic polarization is present.
- (ii) At IR frequency (10^{13} Hz) , ionic and electronic polarization occurs.
- (iii) At audio and radius frequencies (10⁶ to 10¹⁰ Hz), ionic, electronic and orientation polarization occur
- (iv) At very low frequency (10^2 Hz) , all the four polarizations will occur.

47. How does the polarization mechanism in dielectrics depend on temperature?

The ionic and electronic polarization is independent of temperature. In orientation polarization, the thermal agitation will affect the orientation of the polarization. Therefore, the polarization decreases with increase in temperature. In space charge polarization, the increase of temperature increases the migration of ions and hence the polarization increases.

48. What is meant by dielectric loss?

When a dielectric is kept in AC field, the electrical energy is absorbed by material and excess energy is dissipated in the form of heat. This dissipation of energy is called dielectric loss.

49. Define loss angle and loss tangent.

When a voltage is applied across a commercial dielectric, the current lags the voltage by an angle $(90^\circ - \delta)$. The angle, δ , is called loss angle and the tangent value of this angle is called as loss tangent. The loss tangent is a measure of dielectric loss of the material due to resonance.

50. Define dielectric break down.

When a dielectric material is subjected to an electric field, it loses its dielectric property and permits the flow of large amount of current. This phenomenon is called *dielectric breakdown*.

51. Define dielectric strength.

Dielectric strength is defined as the ratio of dielectric breakdown voltage to thickness of the material. Therefore,

 $Dielectric strength = \frac{Dielectric breakdown voltage}{Thickness of the dielectric material}$

52. Mention different types of dielectric breakdown mechanisms.

The different types of dielectric breakdown mechanisms are:

(i) Intrinsic breakdown

- (ii) Thermal breakdown
- (iii) Discharge breakdown
- (iv) Chemical breakdown
- (v) Defects breakdown

53. What is meant by Ferro electricity? Give examples.

Ferro electricity is the phenomenon of spontaneous polarization in dielectric materials. Examples: Rochelle salt, BaTi $\rm O_3$

54. What is Curie temperature?

The temperature at which the Ferro electric behaviour of the material disappears is called Curie temperature.

55. Mention few properties of Ferro electric materials.

The properties of ferro electric materials are:

- (i) They possess spontaneous polarization
- (ii) Ferro electric materials can be easily polarized by very weak magnetic fields.
- (iii) They exhibit hysteresis property
- (iv) They have domain structure similar to ferro magnetic materials.

56. Mention the applications of ferroelectric materials.

The applications of ferroelectric materials are:

- (i) Some ferroelectric materials are used to make posistors. These posistors are used to measure and control temperature like thermistors.
- (ii) The non linear nature of ferroelectric materials can be used to make capacitors with tunable capacitance.
- (iii) These are used in fire sensors and vibration sensors.
- (iv) They are used in high quality infrared cameras.

57. Name a few uses of dielectric materials.

Dielectrics are very widely used as insulating materials as follows:

- (i) Electrical wires are insulated with a outer jacket of plastic.
- (ii) In electric heater, ceramic heads are used in coils to avoid short-circuiting.
- (iii) In electric iron boxes, mica or asbestos insulation is provided to prevent electric shock.
- (iv) In transformers, suitable dielectric material is used for insulating the core and its windings from its container.
- (v) Also in motors and generators, the windings are coated with varnish for insulation.
- (vi) In capacitors, dielectric material is used in between the two electrodes to increase their charge storage capacity.

58. How is dielectric material used in capacitors?

In capacitors, dielectric material is used in between two electrodes to increase the capacitance and charge storage capacity. It is given by

$$C = \frac{\varepsilon_0 \varepsilon_r A}{d}$$

where ε_r is the dielectric constant of the material.

Capacitance with dielectric material = $\varepsilon_r \times$ capacitance without the dielectric material.

59. The dielectric constant of water is **80.** Is water a good dielectric? Is it useful for energy storage in capacitors? Justify your answer.

A dielectric is an insulating material that can be polarized. Water has polar molecules in its purest form as it is a good dielectric fluid. It is not a suitable dielectric material for energy storage in capacitors on account of its ability to conduct even for slight amount of impurity.

60. Define dielectric constant.

The dielectric characteristics of a material are determined by the dielectric constant or relative permittivity ε_r of the material. The dielectric constant is the ratio of the permittivity of the medium (ε) to the permittivity of the space (ε_0). It is given by $\varepsilon_r = \varepsilon/\varepsilon_0$

If the medium has high dielectric constant, it is easily polarized and behaves as a good electrical insulator.

61. Define polarizability and polarization vector.

Polarizability α_e is defined as the ratio of the electric dipole moment p_e to the applied electric field *E*. It is given by $(\alpha_e = p_e/E)$.

Polarization vector \overline{P} is defined as the dipole moment p_e per unit volume V of the dielectric material. Its magnitude is given by $P = p_e/V$.

62. Give the relation between electronic polarization and dielectric constant.

The induced electric dipole moment is proportional to the applied field. This dipole moment per unit volume is called electronic polarization. This is independent of temperature.

The relation between electronic polarization \overline{P} and dielectric constant ε_r is given by

$$\overline{P} = \varepsilon_0 (\varepsilon_r - 1) \overline{E}$$

63. Define electric flux density

The electric flux density or electric induction \overline{D} is defined as the product of applied electric field \overline{E} and the permittivity of the medium ε . It is given by

$$\overline{D} = \varepsilon \overline{E} = \varepsilon_0 \varepsilon_r \overline{E}$$

64. What is electric susceptibility?

The polarization vector \overline{P} is proportional to the total electric field intensity \overline{E} and in the same direction of \overline{E} . It is given by

$$\overline{P} = \varepsilon_0 \chi_e \overline{E}$$

where the constant χ_e is the electric susceptibility. Also, $\chi_e = (\varepsilon_r - 1)$

65. Define polarizability of a dielectric. Give the expression for orientation polarization.

The process of producing electric dipoles which are oriented along the field direction is called polarization in dielectrics. Polarizability of dielectrics α_0 can be defined as induced dipole moment p_0 per unit electric field *E*. It is given by

$$\alpha_o = \frac{p_0}{E}$$

66. Expression for orientation polarization.

The expression for orientation polarization \overline{P}_0 is given by

$$\overline{P}_0 = N\alpha_0\overline{E}$$

where N is the number of molecules per unit volume in dielectrics and \overline{E} is the strength of the applied electric field.

67. Write a note on Clausius-Mosotti relation.

When an elemental dielectric having cubic structure is placed inside an electric field, electronic polarization occurs. If *N* is the number of molecules or atoms per unit volume, the electronic polarizability is related to the dielectric constant (or relative permittivity ε_r) by Clasusus-Mosotti relation.

i.e.,

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{N\alpha_e}{3\varepsilon_0}$$

where ε_0 is the permittivity of free space.

68. Explain electrochemical breakdown in dielectrics.

Chemical and electrochemical type of breakdown is very much related to thermal breakdown. Temperature rise in an insulating material accelerates the chemical deterioration. When temperature rises, the mobility of ions increases and hence, electrochemical reaction takes place. When ionic mobility increases, leakage current also increases and this may lead to dielectric breakdown. Due to field induced chemical reaction, the insulation resistance gradually decreases and finally results in breakdown.

69. Explain intrinsic or Avalanche breakdown.

In a dielectric, the charge displacement increases with increasing field strength. Beyond a critical value of the field strength, there is an electric breakdown due to the physical deterioration of the dielectric material. When the applied electric field is large, some of the electrons in the valence band move towards the conduction band giving rise to large conduction currents. Under this condition, the strength of the local field is of the order of 1 MV/cm. Thus, the liberation or movement of electrons from valence band is called internal field emission of electrons and this breakdown is called the intrinsic breakdown or Avalanche breakdown.
70. Distinguish Lorentz and coulomb forces in dielectrics.

Lorentz force: It is the repulsive force between the nucleus and the electron cloud of a dielectric material when kept in external electric field.

Coulomb force: It is the attractive force between the nucleus and the electron cloud.

71. What is thermal breakdown?

This is due to the attainment of excessive temperature in the dielectric. The electrical energy loss has to be dissipated as heat and if the heat dissipated is less than the heat generated, there is a progressive increase in the temperature of the dielectric which eventually results in local melting. During that time, enormous current will flow through the material and immediately dielectric breakdown will occur.

72. What is discharge breakdown in dielectrics?

The breakdown occurs when the dielectric contains occluded gas bubbles. Since gases require small ionization potential than solids, the gaseous atoms ionize first and the gaseous ions bombard in the solid dielectric causing electrical deterioration.

73. Define relaxation time of a dielectric material.

The relaxation time is the time required for the dipole to reach the equilibrium position from the disturbed position in the presence of an external AC field.

74. What is the effect of frequency of AC electric field on polarization?

As the frequency of applied electric field increases, the total polarization in the material decreases. But at high frequency electric field space charge polarization and orientation polarization do not occur. In DC electric field, all polarization processes occur.

75. Distinguish between dielectric loss and dielectric breakdown.

When a dielectric material is subjected to the AC voltage, the electrical energy is absorbed by the material and is dissipated in the form of heat. This dissipation of energy is called dielectric loss. When the strength of the applied DC electric field exceeds a critical value, the dielectric breakdown occurs allowing a large current to flow through it.

REVIEW QUESTIONS

- 1. Discuss the sources of magnetic dipole moment of the atom.
- 2. Explain Bohr magneton.
- 3. Describe about magnetization.
- 4. Define magnetic permeability and magnetic susceptibility and write the relation between them.
- A magnetic material has a magnetization of 3000 ampere/meter and flux density of 0.044 Wb/m². Calculate the magnetizing force and the relative permeability of the material. (Answer: 203 A · m⁻¹, 17.26)
- A piece of ferric oxide with magnetic field intensity 10⁶ A/m has susceptibility of 1.5 × 10⁻³. Find the magnetization of the material.
 (Answer: 1500 A/m)
- 7. Calculate the magnetic moment per unit volume in copper when subjected to a magnetic field of 10^4 A/m. The magnetic susceptibility of copper is -0.5×10^{-5} . (Answer: -0.05A/m)

3.80 Physics for Electronics Engineering and Information Science

- 8. The saturation value of magnetization of iron is 1.76 × 10⁶ A/m. Iron has body centered cube structure with elementary cube edge of 0.287 nm. Calculate the average number of Bohr magneton contributed to magnetization per atom. (Answer: 2.22 Bohr magnetons)
- g. The paramagnetic material is subjected to a homogeneous field of 10⁶ A/m at room temperature of 30°C. Find the average magnetic moment along direction per spin in Bohr magneton. (Answer: 2.79 × 10⁻³ A · m²)
- 10. Give the classification of magnetic materials and explain their properties.
- 11. Explain ferromagnetic hysteresis.
- 12. Distinguish between soft and hard magnetic materials.
- 13. What are ferrites? Describe the different types of ferrite structure with suitable diagram. What are the applications of ferrites?
- 14. Distinguish between ferromagnetic and antiferromagnetic materials. Give examples.
- 15. Distinguish between dia, para and ferromagnetic materials.
- 16. Compare and contrast ferromagnetism with ferrimagnetism.
- 17. Compare the different types of magnetic materials.
- 18. Discuss about saturation magnetization and Curie temperature.
- 19. Explain in detail the domain theory of ferromagnetism and the types of energy involved in the process of domain growth.
- 20. Give the classification of magnetic materials on the basis of magnetic susceptibility.
- 21. Explain how Bloch wall is formed.
- 22. Describe the ferromagnetic domain theory in detail and how will you account hysteresis of ferromagnetic material.
- 23. What are reversible and irreversible domains? Based on that, explain the phenomenon of hysteresis in ferromagnetic materials.
- 24. Describe the structure of ferrites and mention the properties and applications.
- 25. Explain hysteresis on the basis of domain theory of ferromagnetism.
- 26. Discuss the phenomenon of antiferromagnetism.
- 27. What are hard and soft magnetic materials? Compare their properties and mention their applications.
- 28. Explain how magnetic principle is used in computer data storage.
- 29. Discuss in detail about magnetic hard disk or GMR sensor.
- 30. Write notes on spintronics.
- 31. Describe how read and write process is performed in magnetic hard disk.
- 32. Draw and explain the energy band diagram of dielectric.
- 33. A parallel plate condenser has a capacitance of 6 μ F. The dielectric has permittivity, $\mathcal{E}_r = 100$. For an applied voltage of 3000 V, find the energy stored in the condenser as well as the energy stored in polarizing the dielectric.
- 34. The capacitor is made with a capacitance of 1 μ F working at rms voltage of 50 volt at 1 MHz with a dielectric of thickness 50 μ m. The electric field is 10⁶ volt/m. Given $\mathcal{E} = 3.1 \mathcal{E}_0$ and tan $\delta = 10^3$. Calculate its power loss and power dissipation. (Answer: 172 kW, 15.7 W)
- 35. A parallel plate capacitor of area 650 mm² and a plate separation of 4 mm has a charge of 2 × 10⁻¹⁰ C on it. What is the resultant voltage across the capacitor when a material of dielectric constant 3.5 is introduced between the plates?
 (Answer: 39.73 V)
- 36. What is meant by polarization in dielectrics?
- 37. Define the following: (i) dielectric constant, (ii) polarizability, (iii) polarization vector, (iv) electric flux density and (v) electric susceptibility.
- 38. Write the relation among three electric field vectors.
- 39. Describe the different polarization mechanisms in dielectrics.

- 40. Derive an expression for electronic, ionic and orientational polarizabilities.
- 41. If all the molecular dipoles in a 1 cm radius water drop are pointed in the same direction, calculate the intensity of polarization. The dipole moment of the water molecule is 6×10^{-30} cm. (Answer: 8.4×10^{-10} cm⁻²)
- 42. Calculate the electronic polarizability of argon atom with $\mathcal{E}_r = 1.0024$ and $N = 2.7 \times 10^{25}$ atoms/m³.

(Answer: $7.866 \times 10^{-40} \,\text{F} \cdot \text{m}^2$)

- 43. The atomic weight and density of sulphur are 32 and 2.08 gm/cm³ respectively. The electronic polarizability of the atom is 3.28×10^{-40} F \cdot m². If sulphur solid has cubical symmetry, what will be its relative dielectric constant? (Answer: $\mathcal{E}_r = 2.45$)
- 44. The number of atoms in volume of one cubic metre of hydrogen gas is 9.8 × 10²⁶. The radius of the hydrogen atom is 0.53 Å. Calculate the polarizability and relative permittivity.

(Answer $\mathcal{E}_{e} = 1.655 \times 10^{-41} \,\mathrm{F \cdot m}^{2}$, $\mathcal{E}_{r} = 1.0018$)

- 45. Find the polarizability of CO₂ if its susceptibility is 0.985×10^{-3} . The density of carbon dioxide is 1.977 kg/m^3 . (Answer: $3.24 \times 10^{-40} \text{ F} \cdot \text{m}^2$)
- 46. The radius of the helium atom is about 0.55 Å. Calculate the polarizability of the helium and its relative permittivity. The number of helium atoms in a volume of one metre cube is 2.7 x 10²⁵ atoms.

(Answer: 1.000057)

- 47. NaCl is a cube with lattice parameter 0.564 nm and four formula units per unit cell. The polarizabilities of Na and Cl are 0.16×10^{-40} and 3.30×10^{-40} F·m². Calculate the refractive index of NaCl. (Answer: n = 1.49)
- 48. Determine the percentage of ionic polarization in the sodium chloride crystal which has the optical index of the static dielectric constant 1.5 and 5.6 respectively. (Answer: 51.4 %)
- 49. Derive the Lorentz relation for local field.
- 50. What is meant by local field in a dielectric and how is it calculated for a cubic structure? Hence, deduce Clausius-Mossotti equation.
- 51. The electronic polarizability of Ar atom is 1.75×10^{-40} F·m². Calculate the dielectric constant of solid Ar, using Clausius-Mossotti equation. The density of Ar is given by 1.8×10^3 kg/m³. (Given atomic weight 39.95) (Answer: 1.6536)
- 52. Discuss in detail about dielectric loss. Explain the effect of temperature and frequency on polarization.
- 53. What is meant by loss tangent in dielectrics?
- 54. Derive the dielectric power loss.
- 55. Give a detailed discussion on the various types of dielectric breakdown mechanisms in dielectric materials. What are the remedies to avoid breakdown mechanism?
- 56. What is ferroelectricity? Explain the hysteresis curve exhibited by a ferroelectric material with a suitable sketch. Give examples for ferroelectric materials.
- 57. Write in detail about the application of ferroelectric materials.
- 58. Explain the effects of frequency and temperature on polarization of dielectrics.
- 59. Explain electronic polarization in atoms. Obtain an expression for electronic polarizability in terms of the radius of atoms.
- 60. Explain ionic polarizability in ionic solids and obtain an expression for ionic polarizability in terms of reduced mass of ionic molecule.
- 61. Write short notes on high-*k* dielectrics.

4

Optical Properties of Materials

4.1 INTRODUCTION

Engineering materials play a significant role in our daily life because of their versatile structural properties. The primary physical properties of materials include electrical properties, dielectric properties, magnetic properties, thermal properties and optical properties. Optical properties of materials had a huge impact on the infrastructure development of the communications and the information technology field in recent years. They are also applicable in fields like astronomy, medicine, manufacturing, etc. Optical phenomena are used in devices such as mobile phones, televisions, microscopes, solar cells, etc.

Optical property of any material refers mainly to the interaction of that material with visible light in the electromagnetic spectrum. This chapter first discusses some of the basic principles and concepts related to the nature of light and electromagnetic radiation, and its possible interaction with materials. Then, it explains the optical properties of metals, semiconductors and insulators in terms of absorption, emission and scattering characteristics. Finally, the applications of optical phenomena in photodiode, photodetector, solar cell, light emitting diode (LED), organic LED (OLED) and laser diode are also discussed in this chapter.

4.2 LIGHT AND ELECTROMAGNETIC RADIATION

Electromagnetic radiation is a wave comprising of time varying electric E_x and magnetic field B_y components that are perpendicular to each other and also to the direction of propagation i.e., *z*-direction as shown in Fig. 4.1. In general, visible light, radio waves, X-rays, γ -rays and cosmic rays form part of electromagnetic radiation. The characteristic of each wave is specified by its wavelength λ , and also based on the technique by which it is generated. The electromagnetic spectrum shown in Fig. 4.2 ranges from cosmic rays with wavelengths in the order of 10^{-5} nm, through X-rays, ultraviolet, visible, infrared, microwaves and finally to radio waves with wavelengths up to 10^{-5} m.

The wavelength range of different colours in the visible spectrum is also shown in Fig. 4.2. Here, the visible light is one form of electromagnetic radiation in a narrow region of the spectrum, with wavelengths ranging from 0.4 μ m to 0.7 μ m (400 nm to 700 nm). The colour formation in visible light is determined by its wavelength and white light is a mixture of all colours.



Fig. 4.2 Electromagnetic spectrum

All electromagnetic radiation travels through free space with the velocity of light, $c = 3 \times 10^8$ m/s. This velocity, *c*, is related to electric permittivity of free space, ε_0 and magnetic permeability of free space, μ_0 as

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} \tag{4.1}$$

The above equation gives the relation between the light velocity or electromagnetic constant c and the electrical and magnetic constants. Also, the frequency v and the wavelength λ of the electromagnetic radiation are related to velocity by

$$c = \lambda v \tag{4.2}$$

The refractive index n of the medium is defined as the ratio of the velocity of light in free space c to the velocity in a medium v.

Optical Properties of Materials 4.3

i.e.,
$$n = \frac{c}{v}$$
(4.3)

The velocity of light in a medium is given by

$$v = \frac{1}{\sqrt{\mu\varepsilon}} \tag{4.4}$$

where μ is the magnetic permeability and ε is the electric permittivity of the medium.

Substituting Eqn. (4.1) and Eqn. (4.4) in Eqn. (4.3), we get

$$n = \frac{c}{v} = \frac{1/\sqrt{\mu_0 \varepsilon_0}}{1/\sqrt{\mu \varepsilon}} = \sqrt{\mu_r \varepsilon_r}$$
(4.5)

Here, $\mu = \mu_r \mu_0$ and $\varepsilon = \varepsilon_r \varepsilon_0$ in which μ_r is the relative magnetic permeability and ε_r is the relative electric permittivity (dielectric constant), respectively. Since $\mu_r = 1$ for most of the materials, the refractive index can be written as

$$n \cong \sqrt{\varepsilon_r} \tag{4.6}$$

In quantum-mechanical perspective, the electromagnetic radiation can be considered as groups or packets of energy, rather than as waves. These energy packets or particles are called *photons*. A photon is a massless and chargeless particle. The photon energy E is said to be quantized or can only have specific values as defined by

$$E = hv = \frac{hc}{\lambda} \tag{4.7}$$

where *h* is the Planck's constant, which has the value of 6.626×10^{-34} J·s. The above equation shows that the photon energy is proportional to the radiation frequency or inversely proportional to the wavelength. The optical phenomena such as photoelectric effect, Compton effect etc., can be best understood if light is treated in terms of photons as it involves interaction of radiation with matter. White light is a mixture of seven different colours with its wavelength ranging from 700 to 400 nm and its corresponding energy bandgap from 1.8 to 3.1 eV. Here, the lower energy range corresponds to bandgap of red photons (1.8 eV) and the upper range corresponds to bandgap of blue and violet photons (3.1 eV). Hence, visible light or white light is an electromagnetic radiation with its frequency band ranging from 430 to 750 THz.

EXAMPLE 4.1

The visible light having a wavelength of 500 nm appears green. Compute the frequency and energy of a photon of this light.

Solution Given $\lambda = 500 \text{ nm} = 500 \times 10^{-9} \text{ m}.$

Frequency of a photon of green light,

$$v = \frac{c}{\lambda} = \frac{3 \times 10^8}{500 \times 10^{-9}} = 6 \times 10^{14} \text{ Hz}$$

Energy of a photon of green light,

$$E = hv = 6.626 \times 10^{-34} \times 6 \times 10^{14} = 3.98 \times 10^{-19} \text{ J or } 2.48 \text{ eV}$$

EXAMPLE 4.2

Determine the velocity of light in diamond, which has a dielectric constant ε_r of 5.5 at frequencies within the visible range and a magnetic susceptibility of -2.17×10^{-5} .

Solution Given $\varepsilon_r = 5.5$ and $\chi_m = -2.17 \times 10^{-5}$.

We know that $\mu_r = 1 + \chi_m = 1 - 2.17 \times 10^{-5} \approx 1$

Velocity of light in diamond, $v = \frac{1}{\sqrt{\mu\varepsilon}} = \frac{c}{\sqrt{\mu_r\varepsilon_r}} = \frac{3 \times 10^8}{\sqrt{5.5 \times 1}} = 1.28 \times 10^8 \text{ m/s}$

EXAMPLE 4.3

Zinc selenide (ZnSe) has a band gap of 2.58 eV. Over what range of wavelengths of visible light is it transparent?

Solution Given E = 2.58 eV or $4.13 \times 10^{-19} \text{ J}$. Here, $1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$.

Since the energy bandgap of visible light ranges from 1.8 to 3.1 eV, their corresponding wavelength will range from 700 to 400 nm. The minimum photon energy for visible light is 1.8 eV, which corresponds to a wavelength of 700 nm.

We know that

$$E = hv = \frac{hc}{\lambda}$$

Hence,

$$\lambda = \frac{hc}{E} = \frac{6.626 \times 10^{-34} \times 3 \times 10^8}{4.13 \times 10^{-19}} = 4.81 \times 10^{-7} \approx 480 \,\mathrm{nm}$$

So, ZnSe is transparent to visible light with wavelengths ranging between 480 and 700 nm.

4.3 SPECTRAL RESPONSE OF HUMAN EYE

The human eye is somewhat like a filter with a response similar to that of a tuned circuit. In Fig. 4.3, the solid curve shows the response of an average human eye i.e., photopic vision for normal light levels to the visible part of the spectrum extending from 380 to 760 nm. It is seen that the peak sensitivity occurs at 555 nm and tapers off to zero at 380 and 760 nm. Different colours corresponding to different wavelengths have also been shown. The human eye is more sensitive to green-yellow and less sensitive to violet (shorter wavelength) and red (longer wavelength). The broken curve in Fig. 4.3 with a peak at 507 nm is termed the scotopic eye response for low light levels. The existence of the two responses arises out of the fact that the eye's spectral response shifts at very low light levels.

The lens system of the eye focuses the optical image of the object being observed on the retina. Retina contains light sensitive cellular structures of *rods* and *cones*. The rods sense the brightness levels and the cones are mainly responsible for colour perception. It is estimated that there are some 6,500,000 cones and about 100,000,000 rods connected to the brain through about 800,000 optic nerve fibres.

Horizontal resolution is the ability to resolve horizontal details, i.e., change in brightness levels. Such changes represent vertical edges in the illumination level with frequency determined by the rate of brightness levels. But the human eye follows the abrupt changes in brightness level up to 5 MHz. But for colour images, the eye

can sense the different colours up to 1.5 MHz. Above 1.5 MHz, eye receive only the illumination changes in the object; colours received as black, white and gray.



4.4 LIGHT INTERACTION WITH MATERIALS

The interaction of light i.e., photons with the material gives rise to number of optical phenomena. When light travels from one medium into another e.g., from air into any material, some portion will be transmitted through the medium, some will be absorbed, and some will be reflected at the interface between the two media. The total intensity I_0 of the incident light striking on the surface of any material is the sum of the intensities of the transmitted, absorbed, and reflected light, given by I_T , I_A , and I_R , respectively.

i.e.,

$$I_0 = I_T + I_A + I_R$$

(4.8)

Light can be emitted from a solid when it is stimulated by the source. This phenomenon is called *luminescence*. If the incident energy is in the form of photons, then it is called *photo-luminescence*. If the radiation is produced by the application of an electric field, it is termed as *electro-luminescence*. In *cathode-luminescence*, a beam of electrons bombard the solid and produce radiation of light.

4.5 CLASSIFICATION OF OPTICAL MATERIALS

All materials interact with the incident light rays or photons. These interactions depend on the nature of the material and also on the wavelength of visible light. The interaction of photons with any material is mainly by the combination of absorption, reflection and transmission. The interaction of light with the particles in any material takes place through electronic polarizations and also due to excitation of electrons between different energy states. When the energy gap of material is low compared to visible light, an electron may be excited

by the energy of visible light. Figure 4.4 shows the transmission of light rays through optical materials. Based on the interaction with visible light, optical materials are classified into the following categories:

- (i) Transparent materials
- (ii) Translucent materials
- (iii) Opaque materials

4.5.1 Transparent Materials

Materials that transmit light with minimum absorption and reflection are known as *transparent materials* i.e., objects are easily seen through such transparent optical materials. The light passes through easily if the material is transparent as shown in Fig. 4.4(a). For a material to be transparent, light should not interact with it i.e., the material bandgap energy (E_G) should be larger than the photon energy of visible light ($E_{visible} \approx 1.8-3.1 \text{ eV}$). For example, Sapphire is an insulator with a large bandgap (10 eV) across which it is difficult for electrons to jump. All colored photons are transmitted, with no absorption. Hence, white light is transmitted and the material is said to be transparent. Also, materials such as glass, clean water etc. transmitting much of the incident light with little reflection are called optically transparent. Most of the liquids are highly transparent and responsible for perfect optical transmission due to their pure molecular structure and the absence of structural defects like voids, cracks etc. in it.

4.5.2 Translucent Materials

Materials in which light is transmitted diffusely are known as *translucent materials* i.e. objects are not clearly distinguishable when they are seen through such translucent optical materials. In translucent materials, the incident light is scattered and the diffracted light is allowed on the other side. Here, the light is blurred if the material is translucent as shown in Fig. 4.4(b). Materials such as frosted glass, stained glass, wax, a lampshade, sunglasses, notebook paper, wax paper, tracing paper and vegetable oil form a variety of translucent materials.

4.5.3 Opaque Materials

Materials that do not allow the transmission of light through them are known as *opaque materials* i.e., these materials either absorb or reflect the entire light rays or photons and hence, become opaque. Here, the incident light is blocked if the material is opaque as shown in Fig. 4.4(c). An opaque object is neither transparent nor translucent. Materials such as metals, concrete, wood and stone are opaque to light. Semiconductors such as silicon ($E_G = 1.1 \text{ eV}$), germanium ($E_G = 0.66 \text{ eV}$) and gallium arsenide ($E_G = 1.42 \text{ eV}$) have bandgap energies below the lower bandgap energy range of visible light i.e., 1.8 eV. Hence, photon energy of visible light is greater than the bandgap energy of these semiconductors and is therefore absorbed by these semiconductors. So, these semiconductors are called opaque materials.



Fig. 4.4 Transmission of light through optical materials (a) Transparent (b) Translucent and (c) Opaque

(4.9)

4.6 CARRIER GENERATION AND RECOMBINATION PROCESS

The density of mobile carriers varies due to change in temperature or light shining on a semiconductor. Here, the electron starts out in the valence band, followed by jumping to the conduction band and falling into a trap etc. On a microscopic level, generation-recombination processes occur in a material which leads to the electrons jumping between valence band, conduction band and trap states. This phenomenon is illustrated in Fig. 4.5. The excitation of electrons from the valence band to the conduction band at equilibrium is mainly due to the thermal energy. This form of generation process is referred to as *thermal generation*. Also, a buildup of free carriers is seen when electrons go through continuous excitement from the valence band to the conduction band. This leads to the requirement of a carrier recombination to enable reaching an equilibrium concentration.

Under steady state conditions, the carrier generation and recombination rates are equal.

G = R

i.e.,

The process of carrier generation and recombination in the energy band is shown in Fig. 4.5(a)-(d). Here, the generation of free carriers takes place due to the departure of an electron from the valence band to the conduction band. Such a generation may occur if electrons leave a donor and go into the conduction band. An electron moving from the valence band to an acceptor leads to the generation of holes. Absorption and emission of light is one of the important mechanisms in carrier generation and recombination.



Fig. 4.5 Carrier generation-recombination process in the energy band

4.7 ABSORPTION, EMISSION AND SCATTERING OF LIGHT IN METALS, INSULATORS AND SEMICONDUCTORS

When light interacts with matter, its appearance and colour vary depending on the wavelength of incident light, type of material and the interaction process. This can result in absorption/emission, reflection/scattering and transmission.

Human eye can detect light that has wavelengths in the visible range from 400 nm to 700 nm as shown in Fig. 4.6. Each colour is detected with different efficiency.



Depending on this perception of human eye, materials are classified as opaque, transparent or translucent. Normally, high conductivity materials (metals) have a metallic luster and are opaque. High resistivity materials (insulators) are transparent or translucent and semiconductors can be opaque or transparent with a colour that depends on the band gap of the material.

4.7.1 Optical Properties of Metals

In metals, the valence and conduction bands overlap and hence, all the frequencies of incident radiation from radio waves, through infrared, visible, and into about the middle of the ultraviolet radiation are absorbed by the electrons and excited to states above the Fermi energy. Hence, they are opaque in this region. But most of the radiation is re-emitted or reflected back as the electrons make transitions to their initial state. As the reflectivity for most metals is between 0.90 and 0.95, they possess lustre. Some examples are silver and aluminium. In case of copper and gold, some of the absorbed radiation is not re-emitted and consequently, they appear red-orange and yellow, respectively. But metals are transparent to high frequency radiations such as X-rays and γ -rays.

4.7.2 Optical Properties of Insulators

Insulator like glass is generally transparent to visible light. Therefore, apart from absorption and reflection, *refraction and transmission* have to be taken into account.

Refraction

It is known that when light is incident on a transparent material like a prism, the path of the light ray is deviated with a reduction in its velocity. This is known as *refraction*. The refractive index *n* is defined as the ratio of the velocity of light in air to the velocity of light in the medium. The incident electromagnetic radiation at optical frequencies induces dipoles in the insulator as a result of electronic polarization. These dipoles oscillate in accordance with the different frequency components in the incident radiation. These oscillating dipoles re-radiate the energy at the output end. This causes splitting of the incident radiation into different colours, a reduction in its velocity through the medium and a resultant deviation or bending in its path. Thus, the refractive index *n* is related to the dielectric constant ε_r of the material as $n \cong \sqrt{\varepsilon_r}$.

Generally, the larger an atom is, the greater will be the electronic polarization causing more retardation of the incident electromagnetic radiation. Thus, the velocity will be slower and refractive index will be higher. The refractive indices for several polymers, transparent ceramics, and glasses are listed in Table 4.1.

Materials	Average refractive index	
Polymers		
Polystyrene	1.60	
Polyethylene	1.51	
Polypropylene	1.49	
Poly(methyl methacrylate)	1.49	
Polytetrafluoroethylene	1.35	
Ceramics		
Corundum (Al ₂ O ₃)	1.76	
Periclase (MgO)	1.74	
Spinel (MgAl ₂ O ₄)	1.72	
Glasses		
Dense optical flint glass	1.65	
Quartz (SiO ₂)	1.55	
Soda-lime glass	1.51	
Borosilicate (Pyrex) glass	1.47	
Silica glass	1.458	

Table 4.1 Refractive indices of materials

Transmission

Now, the transparent nature of glass can be analyzed. Glass is an insulator since it has a huge band gap of $E_G \gg 5$ eV. So, it is difficult for electrons to jump across a big energy gap as $E_G \gg E_{G(visible)}$. The bandgap of visible light $E_{G(visible)}$ is from 1.8 to 3.1 eV. Hence, all frequencies will be transmitted with no absorption and the material is transparent. The transmission and absorption is defined by the well-known *Lambert's Law* which states that the absorbance of the material is directly proportional to its thickness (path length).

i.e.,
$$I = I_0 e^{-\alpha x}$$
(4.10)

where I_0 is the incident beam intensity, I is the transmitted beam intensity, x is the distance of light penetration into material from a surface and α is the total linear absorption coefficient in m⁻¹. Here, α takes into account the loss of intensity due to scattering and absorption but it approaches zero for a pure insulator and hence, $I = I_0$.

Reflection/Scattering

Apart from refraction and transmission, *reflection or scattering* can also occur at the interface between two media with different refractive indices. The scattering losses which largely depend on the wavelength of the incident radiation may be due to

- (i) Rayleigh scattering in the low wavelength (UV region) is inversely proportional to wavelength. Hence, longer wavelengths are transmitted.
- (ii) Tyndall effect due to which blue is scattered more than red and hence, longer wavelengths are transmitted.
- (iii) Compton scattering in the X-ray and gamma ray region causing longer wavelengths to be scattered.

Depending on the above, the reflectivity is affected and can be defined as

$$R = \frac{I_R}{I_0} \tag{4.11}$$

where I_0 and I_R are the intensities of the incident and reflected beams, respectively. Reflection losses of optical instruments like lenses can be reduced by coating the reflecting surface with very thin layers of dielectric materials such as magnesium fluoride (MgF₂).

Opacity and Translucency in Insulators

In certain dielectric materials, internal reflection or refraction can also take place at grain boundaries due to the anisotropic variation of refractive index. A transmitted light beam is then deflected and appears diffuse as a result of multiple scattering events. Such materials appear translucent. On the other hand, materials appear opaque when scattering is extensive and hence, not transmitted.

Figure 4.7 shows the difference in optical transmission characteristics in transparent, translucent and opaque crystalline materials. The extent of crystallinity determines the optical property of the material. Scattering is high for highly crystalline materials and hence, they appear translucent or even opaque whereas amorphous materials are fully transparent due to complete transmission.



Fig. 4.7 Variation of light transmittance in crystalline materials (insulators)

4.7.3 Optical Properties of Semiconductors

Semiconductors show strong absorption and reflection in the infrared (IR) region due to interaction of light with optical phonons. Compound semiconductors like GaAs, GaP etc., have a partial ionic character to their bond and exhibit absorption and reflection in the IR region.

Absorption and Emission

In quantum mechanics, electromagnetic radiation is due to particles called photons and each photon carrying energy, hv. Semiconductor devices make the particle nature of EM waves apparent. When a light shining on

a semiconductor can cause an electron from valence band move into the conduction band, electron-hole pairs are generated by this process. Recombination of an electron and a hole followed by emission of light is also possible. Band-to-band transition is the most significant optoelectronic interaction in semiconductors when devices are concerned.



Fig. 4.8 Absorption process in semiconductors from band to band (a) Absorption and (b) Emission

The absorption process from band to band in semiconductors is shown in Fig. 4.8. Here, a photon scatters an electron from the valence band and makes the electron move to the conduction band. Further, in the reverse process, a photon is generated due to recombination of an electron in the conduction band and a hole in the valence band. This leads to emission process. Both absorption and emission processes are employed in light detection and light emission devices. These processes are controlled by laws of conservation of energy and momentum.

Conservation of energy The initial and final energies of the electrons E_i and E_f are seen in the absorption and emission process.

$$E_f = E_i + \hbar \omega$$
 (for absorption) (4.12)

$$E_f = E_i - \hbar \omega$$
 (for emission) (4.13)

where $\hbar\omega$ is the photon energy. This photon energy must be higher than the band gap. The minimum energy difference between the conduction and valence bands is called band gap.

Conservation of momentum Apart from energy conservation, there is also need for conservation of effective momentum $\hbar k$ for the electrons and the photons. The photon k_{ph} value is given by

$$k_{\rm ph} = \frac{2\pi}{\lambda} \tag{4.14}$$

Here, the *k*-value of photons with energies is equal to the band gaps of semiconductors which is in the order of 10^{-4} Å, and is zero in comparison with the *k*-values for electrons. Hence, there is an assurance that the initial and final electrons have the same *k*-value. Thus, law of conservation of momentum ensures that only vertical transitions are allowed during absorption and emission in optical processes, as shown in Fig. 4.8.

Apart from electronic polarization, absorption of a photon can also occur when its energy is greater than the band gap energy of the material. If the photon energy is greater than or equal to the bandgap energy i.e., $hv \ge E_G$, then the photons will be absorbed. But if the photon energy is less than the bandgap energy i.e., $hv < E_G$, then photons will be transmitted.

The minimum wavelength for visible light, λ_{\min} , is about 400 nm, and its maximum wavelength λ_{\max} is about 700 nm. Since $c = 3 \times 10^8$ m/s and $h = 4.13 \times 10^{-15}$ eV s, the maximum possible band gap energy $E_{G(\max)}$ for which absorption of visible light is

$$E_{G(\text{max})} = hv = \frac{hc}{\lambda_{\text{min}}} = \frac{4.13 \times 10^{-15} \times 3 \times 10^8}{400 \times 10^{-9}} = 3.1 \text{eV}$$
$$E_{G(\text{min})} = hv = \frac{hc}{\lambda_{\text{max}}} = \frac{4.13 \times 10^{-15} \times 3 \times 10^8}{700 \times 10^{-9}} = 1.8 \text{eV}$$

Similarly,

Depending on the magnitude of E_G , every non-metallic material becomes opaque at some wavelength. For example, diamond has a band gap of 5.6 eV, and is opaque to radiation having wavelengths less than about 220 nm. Also, semiconducting materials that have band gap energies less than 1.8 eV will absorb all visible light and are opaque. Between 1.8 and 3.1 eV, only partial absorption of the visible spectrum will take place. The materials will appear coloured as shown in Fig. 4.9.



Fig. 4.9 Colours of semiconductors

Consider some examples like shiny black colour of silicon (Si) and yellow colour of gallium phosphide (GaP).

- (i) The energy gap E_G of Si is 1.2 eV. Since $E_{G(visible)}$ is larger than E_G of silicon, all visible light will be absorbed and therefore, silicon appears black. These absorbed photons are scattered by the delocalized electrons in the conduction band. Hence, Si appears shiny.
- (ii) The energy gap E_G of GaP is 2.26 eV and is equivalent to a photon of wavelength, $\lambda = 549$ nm. Here, photons with $E_{G(\text{visible})} > 2.26$ eV (i.e., green, blue, violet) are *absorbed* whereas photons with $E_{G(\text{visible})} < 2.26$ eV (i.e., yellow, orange, red) are *transmitted*. Also, the sensitivity of the human eye is greater for yellow than for red, and hence, GaP appears Yellow or Orange.

Finally to sum up, all the optical processes that occur when light interacts with matter can be depicted as shown in Fig. 4.10.

If *R* is reflectivity, *A* is the absorptivity, and *T* is the transmissivity of the specimen, then the sum of R, *A*, and *T* is unity. The parameters *R*, *A*, and *T* depend on the incident wavelength. Figure 4.11 shows the demonstration of visible light spectrum for a green glass. Here,

there will be variation in wavelength of the fractions of incident light transmitted, absorbed, and reflected through a green glass. For visible light with wavelength of 0.4 μ m through green glass, the fraction of transmitted, absorbed, and reflected energy are approximately 0.90, 0.05, and 0.05, whereas at 0.55 μ m, the corresponding fractions of energy change to 0.50, 0.48, and 0.02, respectively.



Fig. 4.10 Interaction of light with matter



Fig. 4.11 Variation in visible light energy propagating through green glass

EXAMPLE 4.4

Optical radiation with a power density of $2kW/cm^2$ impinges on GaAs. The photon energy is 2.5 eV and the absorption coefficient is 3×10^3 cm⁻¹. Calculate the carrier generation rate at the surface of the sample. If the electron-hole recombination time is 2 ns, calculate the steady state excess carrier density.

Solution Given power density = 2 kW/cm^2 , absorption coefficient = $3 \times 10^3 \text{ cm}^{-1}$, photon energy = 2.5 eV and recombination time = 2 ns.

Carrier generation rate at the surface is

 $G(0) = \frac{\text{Absorption coefficient} \times \text{Power density}}{\text{Photon energy in joule}}$ $= \frac{3 \times 10^3 \times 2 \times 10^3}{2.5 \times 1.6 \times 10^{-19}} = 1.5 \times 10^{25} / \text{cm}^3 \cdot \text{s}$

Excess carrier density, $\delta_n = \delta_p$ = carrier generation rate × recombination time

$$= 1.5 \times 10^{25} \times 2 \times 10^{-9} = 3 \times 10^{16}$$
 / cm³

4.8 PHOTODETECTORS

Semiconductor devices that are used to detect the presence of photons and convert optical signals into electrical signals are called *photodetectors*. The *PN* junction forms the basis of photodetectors which include the photodiode and the phototransistor. When electrons and holes are produced within the depletion region of a *PN* junction, they will be separated by the electric field and hence, a current will be produced. The conductivity of the semiconducting material increases when excess electrons and holes are generated in the material. This change in conductivity forms the basis for any photoconducting device. These photoconductors form the simplest type of photodetectors.

The working principle of photodetectors is based on the rate of absorption of photons and they will absorb photons only if the energy of incident photons is above a certain minimum threshold. The performance of photodetectors can be determined in terms of *figure of merit*. *Responsivity:* The responsivity of any photodetectors is defined as the ratio of the generated photocurrent (*I*) to the amount of optical power (P_0) incident on the detector. It is given by

$$\Re = \frac{I}{P_0} \tag{4.15}$$

The unit of responsivity is amperes/watt.

Quantum Efficiency: Quantum efficiency is a measure of capability of the photodetector in converting all the incident photons into electron-hole pairs. The number of electrons produced per incident photon is defined as the quantum efficiency, which is usually expressed in percentage as,

$$\eta = \frac{\text{Number of electrons produced}}{\text{Number of incident photons}} (\times 100\%)$$

If I is the photocurrent in the external circuit and P_0 is the incident optical power, then

$$\eta = \frac{I/q}{P_0/h\nu} \tag{4.16}$$

where q is the charge of an electron and v is the frequency of incident photon.

Substituting Eqn. (4.16) in Eqn. (4.15), we get

$$\Re = \frac{I}{P_0} = \frac{\eta q}{hv} = \frac{\eta q}{hc} \lambda$$
(4.17)

From the above equation, it is seen that the responsivity depends on wavelength of the photon λ . In case of an ideal photodetector, $\eta = 1$, \Re is linear with λ .

The devices that resulted from the emission of electrons from a metallic surface with vacuum and gas phototubes and were commonly called *photoelectric* devices. However, modern solid state devices, which include emitters, sensors and couplers are called *optoelectronic* devices or electro-optics. The optoelectronic devices are the products of a technology that combines optics with electronics by making use of the optical phenomena such as luminescence, laser, thermal emission, photoconductivity, photovoltaic effect etc. The optoelectronic devices utilize energy in the visible and infrared regions.

Photodetectors are widely used in optical communication systems. There are five main types of photodetectors. They are

- (i) Photoconductor
- (ii) Photodiode
- (iii) PIN Photodiode
- (iv) Avalanche Photodiode and
- (v) Phototransistor

These photodetectors are discussed in detail in the following sections.

4.9 PHOTOCONDUCTOR

When radiation is incident on a semiconductor, some absorption of light by the material takes place, and its conductivity increases. Figure 4.12 shows the photo excitation in semiconductors. This effect is called photoconductive effect which is described as follows.

In a semiconductor material, the forbidden energy gap (E_G) is expressed by

$$E_G = E_2 - E_1 = hv = \frac{hc}{\lambda} \tag{4.18}$$

The energy of a photon is E = hv, where h is the Planck's constant $(6.626 \times 10^{-34} \text{ J} \cdot \text{s})$ and v is the frequency of the incident light. If frequency v is very low so that $E < E_G$, where E_G is the forbidden band energy between valence band and conduction band, the energy is inadequate to transfer an electron from valence band to conduction band and, hence, light passes through the material with very little absorption. However, if $E \ge E_G$, electrons in the valence band absorb the incident photons and get shifted to the conduction band. Also, the conductivity of a semiconductor material is proportional to the concentration of charge carriers as given by

$$\sigma = (n\mu_n + p\mu_p) \tag{4.19}$$

where *n* is the magnitude of free electron concentration, *p* is the magnitude of hole concentration, σ is the conductivity, μ_n is the mobility of electrons and μ_p is the mobility of holes.



Fig. 4.12 Photo excitation in semiconductors

Thus, electron-hole pairs generated by the incident light in addition to those created thermally increases the conductivity (decreases the resistance) of the material resulting in increase in the current in the external circuit. Hence, such a material is called *photoconductor* or *photoresistor*. Therefore, for photoconduction to take place in an intrinsic semiconductor, the photon must possess energy at least equal to the forbidden energy gap E_G . Thus, the minimum frequency v_c to cause photoconduction is given by

$$v_c = \frac{E_G}{h} \tag{4.20}$$

The long wavelength limit, cut-off wavelength, or the critical wavelength of the material, λ_c , is obtained by substituting the values of *c* and *h* resulting in

$$\lambda_c = \frac{1.24}{E_G(\text{eV})} \mu \text{m} \tag{4.21}$$

For silicon, $E_G = 1.1$ eV and $\lambda_c = 1.13 \ \mu$ m, whereas for germanium, $E_G = 0.72$ eV and $\lambda_c = 1.75 \ \mu$ m at room temperature.

4.9.1 Bulk Type Photoresistor or Photoconductive Cell

The *photoconductive cell* (*PC*) or photodetector is a two-terminal device which is used as a Light Dependent Resistor (LDR). It is made of a thin layer of semiconductor material such as cadmium sulphide (CdS), lead sulphide (PbS), or cadmium selenide (CdSe) whose spectral responses are shown in Fig. 4.13. The photoconducting device with the widest applications is the CdS cell, because it has high dissipation capability, with excellent sensitivity in the visible spectrum and low resistance when stimulated by light. The main drawback of CdS cell is its slower speed of response. PbS has the fastest speed of response.



The illumination characteristics of photoconductive detectors are shown in Fig. 4.14(a). It exhibits the peculiar property that its resistance decreases in the presence of light and increases in the absence of light. The cell simply acts as a conductor whose resistance changes when illuminated. In absolute darkness, the resistance is as high as 2 M Ω and in strong light, the resistance is less than 10 Ω .



Fig. 4.14 Photoconductive detector (a) Illumination characteristics and (b) its circuit

A simple circuit for a photoconductive detector is shown in Fig. 4.14(b). The semiconductor layer is enclosed in a sealed housing. A glass window in the housing permits light to fall on the active material of the cell. Here, the resistance of the photoconductive detector, in series with R, limits the amount of current I in the circuit. The ammeter A is used to measure the current I. When no light falls on the cell, its resistance is very high and the current I is low. Hence, the voltage drop V_o across R is relatively low. When the cell is illuminated, its resistance becomes very low. Hence, current I increases and voltage V_o increases. Thus, this simple circuit arrangement with slight modification can be used in control circuits to control the current.

Applications

The detector is used either as an ON/OFF device to detect the presence or absence of a light source which is used for automatic street lighting or some intermediate resistance value can be used as a trigger level to control relays and motors. Further, it is used to measure a fixed amount of illumination and to record a modulating light intensity.

It is used in counting systems where the objects on a conveyor belt interrupt a light beam to produce a series of pulses which operates a counter.

It is used in twilight switching circuits. When the day light has faded to a given level, the corresponding resistance of the detector causes another circuit to switch ON the required lights.

It is widely used in cameras to control shutter opening during the flash. Twin photoconductive cells mounted in the same package have been used in optical bridge circuits for position control mechanisms and dualchannel remote volume control circuits.

4.10 PHOTOCURRENT IN A PN PHOTODIODE

Silicon photodiode is a light-sensitive device, which converts light signals into electrical signals. The construction and symbol of a photodiode are shown in Fig. 4.15. The diode is made of a semiconductor PN junction kept in a sealed plastic or glass casing. The cover is so designed that the light rays are allowed to fall on one surface across the junction. The remaining sides of the casing are painted to restrict the penetration of light rays. A lens permits light to fall on the junction.



Fig. 4.15 Photodiode (a) circuit and (b) symbol

When light falls on the reverse-biased *PN* photodiode junction, hole-electron pairs are created. The movement of these hole-electron pairs in a properly connected circuit results in current flow. The magnitude of the photocurrent depends on the number of charge carriers generated and hence, on the illumination of the diode element. This current is also affected by the frequency of the light falling on the junction of the photodiode.

The magnitude of the current under large reverse bias is given by

$$I = I_{S} + I_{o}(1 - e^{V/\eta V_{T}})$$
(4.22)

where I_o is the reverse saturation current, I_S is the short-circuit current which is proportional to the light intensity, V is the voltage across the diode, V_T is the volt equivalent of temperature and η is the intrinsic parameter which is 1 for Ge and 2 for Si. The characteristics of a photodiode are shown in Fig. 4.16. The reverse current increases in direct proportion to the level of illumination. Even when no light is applied, there is a minimum reverse leakage current called *dark current*, flowing through the device. Germanium has a higher dark current than silicon, but it also has a higher level of reverse current.

Applications

Photodiodes are used as light detectors, demodulators and encoders. They are also used in optical communication



Fig. 4.16 Characteristics of photodiode

system, high-speed counting and switching circuits. Further, they are used in computer card punching and tapes, light operated switches, sound track films and electronic control circuits.

4.11 PIN PHOTODIODE

PIN photodiode is used for the detection of light at the receiving end in optical communication. It is a threeregion reverse-biased junction diode. A layer of intrinsic silicon is sandwiched between heavily doped *P* and *N*-type semiconductor materials. As shown in Fig. 4.17, the depletion region extends almost to the entire intrinsic layer where most of the absorption of light photons takes place. The width of the intrinsic layer is large compared to the width of the other two layers. This ensures large absorption of light photons in the depletion region which also forms the absorption region. Light photons incident on the *PIN* photodiode are absorbed in the absorption region which leads to the generation of electron-hole pairs. These charge carriers present in the depletion region drift under the influence of the existing electric field that is set up due to the applied reverse bias. The reverse current flowing in the external circuit increases linearly with the level of illumination.



Fig. 4.17 Structure of PIN photodiode

As the process of drifting is quicker than diffusion, the transit time of the charge carriers is small so that the response time is considerably reduced. The large width of the depletion region results in achieving high quantum efficiency.

4.12 AVALANCHE PHOTODIODE (APD)

Avalanche photodiode (APD) is used in optical communication for detection of light at the receiving end. It converts the input light signal into electrical signal. The structure of an APD is shown in Fig. 4.18. It essentially consists of reverse-biased *PN* junction. The depletion region in this reverse-biased *PN* junction is formed by immobile positively charged donor atoms in the *N*-type semiconductor material and immobile negatively charged acceptor atoms in the *P*-type material. The electric field in this depletion region is very high where most of the photons are absorbed and primary charge carriers (electron-hole pair) are generated. These charge carriers acquire sufficient energy from the electric field to excite new electron-hole pairs by a process known as impact ionization. These new carriers created by impact ionization can themselves produce additional carriers by the same mechanism.



Fig. 4.18 Structure of APD

For this process, APD requires a high reverse-bias voltage in the order of 100-400 V. Carrier multiplication factors as great as 10^4 may be obtained using defect free materials. Electron-hole pairs, thus generated, separate and drift under the influence of the electric field in the depletion region and diffuse outside the depletion region so that they are finally collected in the detector terminals. This leads to a flow of current in the external circuit whose magnitude is proportional to the intensity of light incident on APD. Due to the internal gain mechanism in an APD, a large electrical response is obtained even for a weak input light signal. Quantum efficiency closer to 100% in the working region can be obtained.

4.13 PHOTOTRANSISTOR

The phototransistor is a much more sensitive semiconductor photodevice than the *PN* photodiode. The current produced by a photodiode is very low which cannot be directly used in control applications. Therefore, this current should be amplified before applying to control circuits. The phototransistor is a light detector which combines a photodiode and a transistor amplifier. When the phototransistor is illuminated, it permits a larger flow of current.



Fig. 4.19 NPN phototransistor (a) Symbol (b) Biasing arrangement

Figure 4.19 shows the circuit of an *NPN* phototransistor. It is usually connected in a common emitter (CE) configuration with the base open. A lens focuses the light on the base-collector junction. Although the phototransistor has three sections, only two leads, the emitter and collector leads, are generally used. In this device, base current is supplied by the current created by the light falling on the base-collector photodiode junction.

When there is no radiant excitation, the minority carriers are generated thermally, and the electrons crossing from the base to the collector and the holes crossing from the collector to the base constitute the reverse saturation collector current I_{CO} . With $I_B = 0$, the collector current is given by

$$I_C = (1+\beta)I_{CO} \tag{4.23}$$

When the light is turned ON, additional minority carriers are photogenerated and the total collector current is

$$I_{C} = (1 + \beta)(I_{CO} + I_{L})$$
(4.24)

where I_L is the reverse saturation current due to the light and β is the current gain of the phototransistor.

The current in a phototransistor is dependent mainly on the intensity of light entering the lens and is less affected by the voltage applied to the external circuit. Figure 4.20 shows a graph of collector current I_C as a function of collector-emitter voltage V_{CE} and as a function of illumination H.



The phototransistors find extensive applications in high-speed reading of computer punched cards and tapes, light detection systems, light operated switches, reading of film sound track, production line counting of objects which interrupt a light beam, etc.

4.14 PHOTODETECTOR AS OPTOCOUPLER

Photodetectors are light sensing devices that convert optical energy into electrical energy i.e., converting light photons into current. It works on the principle of converting photons to electrons. Photons contribute their energies to electrons in the valence band and excite them to the conduction band. As a result, electron-hole pairs are generated which increases the conductivity of the material. Photodetectors are widely used in optical communication systems. Here, the photodetectors receive the transmitted optical pulses and convert them into electronic pulses with little attenuation. These electronic pulses can be used by telephones, computers or any other terminals in the receiving end. Optoelectronic devices such as photodiodes, PIN photodiodes, avalanche photodiodes and phototransistors form part of photodetectors.

One of the main applications of photodetector is optocoupler, which is useful in optical communications. An optocoupler is a solid-state component in which the light emitter, the light path and the light detector are all enclosed within the component and cannot be changed externally. As the optocoupler provides electrical isolation between circuits, it is also called *optoisolator*. An optoisolator allows signal transfer without coupling wires, capacitors or transformers. It can couple digital (ON/OFF) or analog (continuous) signals.

The schematic representation of photodetector as an optocoupler is shown in Fig. 4.21. The optoisolator, also referred to as an optoelectronic coupler, generally consists of an infrared LED and a photodetector such as PIN photodiode for fast switching, phototransistor Darlington pair, or photo-SCR combined in a single package. Optoisolators transduce input voltage to proportional light intensity by using LEDs. The light is transduced back to output voltage using light sensitive devices. GaAs LEDs are used to provide spectral matching with the silicon sensors.



Fig. 4.21 Schematic representation of photodetector as optocoupler

The wavelength response of each device is made to be as identical as possible to permit the highest measure of coupling possible. There is a transparent insulating cap between each set of elements embedded in the structure (not visible) to permit the passage of light. They are designed with very small response times in such a way that they can be used to transmit data in the MHz range.

The rigid structure of this package permits one-way transfer of the electrical signal from the LED to the photodetector, without any electrical connection between the input and output circuitry. The extent of isolation between input and output depends on the kind of material in the light path and on the distance between the light emitter and the light detector. A significant advantage of the optoisolator is its high isolation resistance of the order of $10^{11} \Omega$ with isolation voltages up to 2500 V between the input and output signals, and this feature allows it to be used as an interface between high voltage and low voltage systems. Application for this device includes the interfacing of different types of logic circuits and their use in level-and-position-sensing circuits.

In the optoisolator, the power dissipation of LED and phototransistor are almost equal and the common emitter reverse saturation current I_{CEO} is measured in nano-amperes. The relative output current is almost constant when the case temperature varies from 25 to 75°C. The V_{CE} voltage affects the resulting collector current only very slightly. The switching time of an optoisolator decreases with increased current, while for many devices it is exactly the reverse. It is only 2 ms for a collector current of 6 mA and a load resistance of 100 Ω . The schematic diagrams of optoisolator with a photodiode, photo-Darlington pair, and photo-silicon controlled rectifier (SCR) are illustrated in Fig. 4.22.



Fig. 4.22 Optoisolators (a) Photodiode (b) Photo-Darlington pair and (c) Photo-SCR

4.15 PHOTOVOLTAIC CELL

If the *PN* junction is open circuited, the light energy is used to create a potential difference which is proportional to the frequency and intensity of the incident light. This phenomenon is called *photovoltaic effect*. The photovoltaic cell, a light-sensitive semiconductor device, produces a voltage when illuminated which may be used directly to supply small amounts of electric power. In the photovoltaic device without any applied voltage, the junction generates a voltage depending upon the illumination and the load. The voltage generated is due to the accumulation of carriers produced by photon excitation.

The photovoltaic potential is the voltage at which zero resultant current is obtained under open-circuited conditions. The photovoltaic emf is 0.5 V for either silicon or selenium cell and 0.1 V for germanium cell the short circuit cell current is of the order of 1 mA. The magnitude of the current under large reverse bias is given by

$$I = I_{S} + I_{a}(1 - e^{V/\eta V_{T}})$$
(4.25)

The photovoltaic voltage V_{max} which corresponds to an open-circuited diode can be obtained by substituting I = 0 in the above equation. Hence,

$$V_{\max} = \eta V_T \ln \left(1 + \frac{I_s}{I_o} \right) \tag{4.26}$$

As $I_S >> I_o$, V_{max} increases logarithmically with short-circuit current, I_S , and hence, with illumination. The voltage increases as the intensity of light falling on the semiconductor junction of this cell increases. A photovoltaic cell consists of a piece of semiconductor material such as silicon, germanium, or selenium which is bonded to a metal plate, as shown in Fig. 4.23(a). The circuit symbol for photovoltaic cell is shown in Fig. 4.23(b).



Fig. 4.23 Photovoltaic cell (a) Construction (b) Circuit symbol

The spectral responses of silicon, germanium and selenium are shown in Fig. 4.24, indicating that photoconductor is a frequency-selective device. As the spectral response of silicon and germanium extends well into infrared region, its efficiency is quite high. The selenium cell has two advantages over silicon, viz., (i) its spectral response is almost similar to that of the human eye, and (ii) it has the ability to withstand damaging radiation environments, lasting up to 10,000 times longer than silicon.



Fig. 4.24 Spectral responses of Si, Ge and Se

The characteristic curves of output voltage versus light intensity and output current versus light intensity are shown in Figs. 4.25(a) and (b), respectively. Photovoltaic cells are used in low-power devices such as light meters. Nowadays, with an improvement in the efficiency of these cells, more power is produced, as in solar cells which are photovoltaic devices. When operated in the short-circuit mode, the current is proportional to the illumination and photovoltaic cell is used to construct a direct-reading foot-candle meter.



Fig. 4.25 Characteristic of photovoltaic cell (a) Output voltage vs. light intensity and (b) Output current vs. light intensity

4.16 SOLAR CELL

When sunlight is incident on a photovoltaic cell, it is converted into electric energy. Such an energy converter is called solar cell or solar battery and is used in satellites to provide the electrical power. This cell consists of a single semiconductor crystal which has been doped with both *P*- and *N*-type impurities, thereby forming a *PN* junction.



Fig. 4.26 Basic construction of a PN junction solar cell

The basic construction of a PN junction solar cell is shown in Fig. 4.26. Sunlight incident on the glass plate G passes through it and reaches the junction. An incident light photon at the junction may collide with a valence electron and impart sufficient energy to make a transition to the conduction band. As a result, an electron-hole pair is formed. The newly formed electrons are minority carriers in the P-region. They move freely across the junction. Similarly, holes formed in the N-region cross the junction in the opposite direction. The flow of these electrons and holes across the junction is in a direction opposite to the conventional forward current in a PN junction. Further, it leads to the accumulation of majority carriers on both sides of the junction. This gives rise to a photovoltaic voltage across the junction in the open-circuit condition. This voltage is a logarithmic function of illumination.

In bright sunlight, about 0.6 V is developed by a single solar cell. The amount of power, the cell can deliver depends on the extent of its active surface. An average cell will produce about 30 mW per square inch of surface, operating in a load of 4 W. To increase the power output, large banks of cells are used in series and parallel combinations. The efficiency of the solar cell is measured by the ratio of electric energy output to the light energy input expressed as a percentage. At present, efficiency in the range of 10 to 40% is obtained. Silicon and selenium are the materials used widely in solar cells because of their excellent temperature characteristics. Solar cell assemblies are used to make solar modules that generate electrical power from sunlight. A solar array generates solar power using solar energy.

4.17 LIGHT EMITTING DIODE (LED)

The Light-Emitting Diode (LED) is a PN junction device which emits light when forward biased, by a phenomenon called electroluminescence. In all semiconductor PN junctions, some of the energy will be radiated as heat and some in the form of photons. In silicon and germanium, greater percentage of energy

is given out in the form of heat and the emitted light is insignificant. In other materials such as gallium phosphide (GaP) or gallium arsenide phosphide (GaAsP), the number of photons of light energy emitted is sufficient to create a visible light source. Here, the charge carrier recombination takes place when electrons from the *N*-side cross the junction and recombine with the holes on the *P*-side.

The LED under forward bias and its symbol are shown in Figs. 4.27(a) and (b), respectively. When an LED is forward biased, the electrons and holes move towards the junction and recombination takes place. As a result of recombination, the electrons lying in the conduction bands of *N*-region fall into the holes lying in the valence band of a *P*-region. The difference of energy between the conduction band and the valence band is radiated in the form of light energy. Each recombination causes radiation of light energy. Light is generated by recombination of electrons and holes whereby their excess energy is transferred to an emitted photon. The brightness of the emitted light is directly proportional to the forward-bias current.

Figure 4.27(c) shows the basic structure of an LED showing recombination of carriers and emission of light. Here, an *N*-type layer is grown on a substrate and a *P*-type is deposited on it by diffusion. Since carrier recombination takes place in the *P*-layer, it is kept uppermost. The metal anode connections are made at the outer edges of the *P*-layer so as to allow more central surface area for the light to escape. LEDs are manufactured with domed lenses in order to reduce the reabsorption problem. A metal (gold) film is applied to the bottom of the substrate for reflecting as much light as possible to the surface of the device and also to provide cathode connection. LEDs are always encased to protect their delicate wires.

The efficiency of light generation increases with increase in injected current and with decrease in temperature. The light is concentrated near the junction as the carriers are available within a diffusion length of the junction.



Fig. 4.27 Light Emitting Diode (a) LED under forward bias (b) Symbol (c) Recombination and emission of light

LEDs radiate different colours such as red, green, yellow, orange, blue and white. Some of the LEDs emit infrared (invisible) light also. The wavelength of emitted light depends on the energy gap of the material. Hence, the colour of the emitted light depends on the type of material used is given as follows:

- (i) Gallium arsenide (GaAs) infrared radiation (invisible)
- (ii) Gallium phosphide (GaP) red or green
- (iii) Gallium arsenide phosphide (GaAsP) red or yellow

In order to protect LEDs, resistance of 1 kW or 1.5 kW must be connected in series with the LED. LEDs emit no light when reverse biased. LEDs operate at voltage levels from 1.5 to 3.3 V, with the current of some tens of milliamperes. The power requirement is typically from 10 to 150 mW with a lifetime of 1,00,000 + hours. LEDs can be switched ON and OFF at a very fast speed of 1 ns.

They are used in burglar alarm systems, picture phones, multimeters, calculators, digital meters, microprocessors, digital computers, electronic telephone exchange, intercoms, electronic panels, digital watches, solid state video displays, and optical communication systems. Also, there are two-lead LED lamps which contain two LEDs, so that a reversal in biasing will change the colour from green to red, or vice-versa.

When the emitted light is coherent, i.e., essentially monochromatic, then such a diode is referred to as an Injection Laser Diode (ILD). The LED and ILD are the two main types used as optical sources. ILD has a shorter rise time than LED, which makes the ILD more suitable for wide-bandwidth and high-data-rate applications. In addition, more optical power can be coupled into a fibre with an ILD, which is important for long distance transmission. A disadvantage of the ILD is the strong temperature dependence of the output characteristic curve.

4.17.1 Infrared Emitters

The infrared emitting diodes are PN junction gallium arsenide devices which emit a beam of light when forward biased. When the junction is energized, electrons from the *N*-region will recombine with the excess holes of the *P*-material in a specially formed recombination region sandwiched between the *P*- and *N*-type materials. This recombination, which tends to restore the equilibrium carrier densities, can result in the emission of photons from the junction. The radiant energy from the device is infrared with a typical peak at 0.9 mm, which ideally matches the response of silicon photodiode and phototransistors.

These infrared emitting diodes are used in shaft encoders, data-transmission systems, intrusion alarms, card and paper tape readers, and high density mounting applications. The shaft encoder can produce 150 mW of radiant energy at 1.2 V and 50 mA.

4.18 ORGANIC LED (OLED)

The technology used in display gradually evolved from the bulky, high power cathode ray tubes (CRT) to the latest light weight, low power plasma displays, LCDs and LEDs. Light-emitting diode in which an organic compound is used as a film in its layer is known as organic light-emitting diode (OLED). The layer of organic semiconductor is located between two electrodes, thereby forming an emissive electroluminescent layer which emits light in response to an electric current. Display using OLED is one of the fascinating substitutes to all the predecessor display technologies. OLEDs are mainly used to make digital displays in devices such as mobile phones, television screens and computer monitors. There are generally two types of OLED displays. They are (i) Passive matrix organic LED (PMOLED) display and (ii) Active matrix organic LED (AMOLED) display.

4.18.1 Passive Matrix Organic LED Display

The PMOLED display is systematically organized in a ribbed type structure comprising of rows and columns, as shown in Fig. 4.28. The cathode rows are superimposed over the anode columns in the display. There is an organic film lying in the middle of the anode and cathode. Due to the event of switching ON any specific row and column lines, each pixel positioned at their intersection points are kindled.



Fig. 4.28 Formation of pixel in PMOLED

Despite being simple in design and fabrication, PMOLED display is complex due to the fact that every line requires limitation in current for every single diode. The arrangement varies with respect to the number of diodes which get activated in the respective rows. The time taken to drive every single pixel decreases with the number of rows. The decrease in time taken claims that the peak brightness to be kept in high levels, thus causing the row line current to increase.

In order to attain a split addressed Video Graphics Array (VGA) display of average luminance 300 nits, the peak brightness level of 72,000 nits was essential. This deciphered 1.6A of peak current level on the row lines. In passive matrix addressing displays, adequate mean brightness is achieved by large driving current. This current is critical due to the involvement of high drive voltages. This causes outsized voltage drops in row lines, which in turn pushes the OLED working to a less power efficient - high voltage rule.

The effects of higher voltages are high power dissipation, more flicker noise and reduced lifetimes. The lifetime is shortened due to the fact that the high level of brightness leads to increase in display driving voltage which automatically trades off with the expected level of lifetime. The association between driving voltage level and lifetime is almost linear in most of the OLED products. Hence, in order to overcome such quandaries, active matrix addressing is employed. PMOLED has large power consumption when compared to AMOLED display.

Advantages

Large scale fabrication at low disbursement is made feasible, as the organic film and cathode in the PMOLED display are deposited by standard processing techniques.

Applications

Due to its high power consumption, PMOLED display restricts its use in applications which satisfies the display size of 50-80 mm as measured diagonally or the number of rows having less than 100. It is mostly used in small applications that uses text and icon displays like cell phones, automobiles etc.

4.18.2 Active Matrix Organic LED Display

The main benefits of employing active matrix in OLED are it requires only low voltage for operation and uses very small peak pixel currents, thus resulting in much better efficiency and brightness. As the brightness

is directly proportional to the current driven via OLED, the main aim of active matrix is to deliver a constant current flow throughout the full frame time, thereby eliminating the high current which is a drawback in passive matrix method.

AMOLED display consists of matrix similar to PMOLED display. Every pixel comprises of an anode, a cathode and an intervening organic film as shown in Fig. 4.29. Here, to visualize the OLED structure which is located underneath, only half the portion of the cathode plate is drawn. The respective scan and data lines are provided for selective pixel addressing.



Fig. 4.29 AMOLED activation through TFT

The exceptional feature of this type of display is that the pixels get activated by thin film transistors (TFT) array. The current flow through each pixel is monitored and controlled by this array and it also causes its discriminatory activation. The light is emitted by the pixel when current flows inside a particular pixel. A pixel is made up of two transistors. The ON and OFF control of pixel is done by one transistor and constant current for the pixel is provided by another transistor. The AMOLED display pixel circuit consists of a switching transistor, a driver transistor and one storage capacitor. The necessary current required for the OLED pixels is supplied by driver transistors. This current is in the range of 1–5 μ A per pixel to display signals of high resolution. Therefore, the requirement of high drive current as needed by PMOLED display is avoided in AMOLED display.

A vital technology which majorly sets the limits of display industry is known as the *backplane technology*. The backplane is responsible for turning ON and OFF of each pixel. It is constructed using thin film transistor (TFT) that acts as a switch. So, it is generally called TFT backplane. The power consumption of

TFT backplane is very low. These get refreshed fast unlike LCD and there is increased convenience for TV and other displays which broadcast moving pictures and graphics.

Advantages

Mostly, the high-performance applications use AMOLED displays, due to its beneficial features of long viewing and broad visibility range coupled with wide viewing angles. It has slimness of about 1.3 mm and has small dot pitch size. It has high level of color saturation and contrast. It also has fast response time in μ s, thus minimizing the blurring effect. It has reduced heat dissipation and low temperature operation.

Applications

The main issue in AMOLED is its less lifetime. The initial products have lifetimes of about 15,000 hours, which definitely need improvement. This limits their usage in TV or computer applications. But these displays are now being used in TV and there are research efforts under process to increase the lifetime. Also, these are mainly used in mobile phones. The low cost of AMOLED displays favours its use in many applications.

4.19 LASER DIODE

Similar to LED, LASER diodes are used to convert the electrical signal to light signal. In direct bandgap materials where high recombination velocities exist, optical gain can be achieved by creating population inversion of carriers through high-level current injection and by forming a resonant cavity. This cavity is usually produced by the high Fresnel reflectivity obtained from cleaving the material along faces perpendicular to the junction plane.



Fig. 4.30 LASER diode (a) structure and (b) characteristics

The structure and characteristics of a typical LASER diode is shown in Fig. 4.30. In this diode, opposite ends of the junction are polished to get mirror like surfaces. When free electrons recombine with holes, the emitted photons reflect back and forth between the mirror surfaces. The region between the mirrored ends acts like a cavity that filters the light and purifies it colour. As the photons bounce back and forth, they induce an avalanche effect that causes all newly created photons to be emitted with the same phase. One of the mirror surfaces is semitransparent. From this surface, a fine threadlike beam of photons emerge out. All the photons of LASER light have same frequency and phase and, hence, coherent.

4.30 Physics for Electronics Engineering and Information Science

It has a well-defined current threshold as seen from the power output vs drive current characteristic. Below this threshold, the device exhibits low levels of spontaneous emission. At the limiting current density, stimulated emission occurs and the emitted radiation increases linearly with drive current. Table 4.2 lists the characteristics of LED and LASER diode.

Characteristics	LED	LASER diode
Structure	Double heterojunction	Double heterojunction
Materials used	GaAs, GaP, GaAsP, InGaAsP, InP, InGaN, AlGaAs, AlGaInP	GaAs, GaN, InGaAsP, InP
Output radiation	Incoherent (spontaneous emission)	Coherent (Stimulated emission)
Fibre type	Multimode only	Both single mode and multimode
Spectral line width, $\Delta \lambda$	Wider i.e., 100 nm	Narrower i.e., 2 – 4 nm for multimode laser and < 0.1 nm for single mode laser
Rise time	5 – 20 ns	< 1 ns
Current	Drive current: 50 to 100 mA peak	Threshold current: 5 to 40 mA
Output Power	Linearly proportional to drive current	Proportional to current above threshold
Ease of operation	Easier	More difficult
Cost	Less	Expensive

 Table 4.2
 Comparison of LED and LASER diode

4.20 EXCITONS

When a photon strikes an optical material, it gets absorbed. Using this photonic energy, the electron can excite itself from the valance band to the conduction band. Figure 4.31(a) shows the energy-momentum (*E-k*) diagram for the given optical materials. Figure 4.31(b) shows the *E-k* diagram for generation of electronhole.

As the electron jumps, it leaves behind a positively charged hole due to the Coulomb's force from large number of electrons surrounding the hole. As a result, this electron in the conduction band experiences a force of attraction towards the hole. The combination of the electron and the hole together, which are attracted to each other by the electrostatic Coulomb force is called an *exciton*. Figure 4.31(c), shows the picture of optical material for two-particles where both the charges are energized. The momentum wave vector for an exciton is

$$k = k_e + k_h$$

As shown in Fig. 4.31(d), $E_c + (\hbar^2 k^2 / 2m)$ is the new energy obtained by optical transition. Generally, energy gap will be high comparing to photonic energy. Unbound electron and hole will possess high energy balance comparing to this attraction, but this attraction is stable.

Exciton has smaller binding energy and huge size. This is almost the size of a hydrogen atom. This bound state of an exciton is called as hydro-genic. Exciton structure is stable because of the smaller effective mass by the excited electron and hole. Spin of the electron and hole will be parallel or anti-parallel. Fine structure

of exciton is framed by exchange interaction due to the spin (gets coupled). The properties of an exciton depend on momentum (k-vector).



Fig. 4.31 Optical absorption and transition of exciton (a) Incident photon on any material being absorbed (b) An electron in n = 1 state on absorbing photon undergoes optical transition to create a hole (c) exciton wave vector for two particle and (d) new energy obtained through optical transition

Exciton is known as an electrically neutral-quasi particle which is usually found in materials like insulators, semiconductors and some liquids that have low dielectric values. Exciton has the capability to transport the energy without possessing the net electric charge. Figure 4.32 shows different states of an exciton. An electron orbiting a proton, i.e., a hole due to Coulomb force is shown in Fig. 4.32(a). Due to attraction, the electron-hole pair is bounded tightly in the same unit cells or in the neighbor unit cells, specifically in an ionic crystal. This is referred to as the Frenkel exciton. Here, one electron is promoted from the valence band to the conduction band, thereby leaving a localized positively charged hole as shown in Fig. 4.32(b). The first order decay reaction showing the radioactive recombination of an exciton resulting in photoluminescence is illustrated in Fig. 4.32(c).



Fig. 4.32 Different states of an exciton (a) Electron orbiting a hole (b) Frenkel exciton (c) first order decay reaction (d) decomposition reaction (e) Release of phonon due to non-radioactive decay (f) second order decay reaction

Figure 4.32(d) shows the decomposition reaction. The release of phonon due to non-radioactive decay is shown in Fig. 4.32(e). The exciton-exciton annihilation (EEA) due to second order decay reaction is shown in Fig. 4.32(f). Here, on collision of the two excitons, one is annihilated whereas the other one uses the energy to be shifted to the higher energy level. Semiconductors that have a large value of dielectric constant possess a weak electron-hole interaction due to valence electrons. Hence, it has weakly bound Wannier-Mott excitons.

4.21 QUANTUM CONFINED STARK EFFECTS (QCSE)

The alignment of fundamental particles on application of external magnetic or electric field can be understood using the atomic structure of atoms and molecules considered in the light of quantum. Shifting and splitting of spectral lines occurs when an external electric field is applied to a specified atom or molecule is known as *stark effect* and the broadcasting of spectral lines are referred as *stark shift*. This stark effect has been applied in communication and electronics industries and also, it has been considered in different optical materials.

Stark spectroscopy was recognized as a tool to realize the structure of atoms and molecules earlier. As a result of the technological advancement in semiconductor industry and also, due to the idea of exciton, it has been considered that bound exciton plays key role to improve the stark effect. A small band gap semiconductor material is inserted between two layers of a large band gap material for designing hetero-structure and it acts as a quantum well. This is a perfect example for quantum confinement. Whenever the electrons and holes are restricted in the discrete energy levels, there will be energy sub-bands, in the state where no electric field is applied. Hence, it is concluded that only some frequencies of light may be absorbed or emitted by this system.

Due to the application of electric field, the electron and hole forms exciton which will be extracted in opposite direction, and it leads to lowering of overlap integral. Further, it will reduce the recombination efficiency of the system, which is called quantum fluorescence yield. Hole states are shifted to higher energies at the same time as the electron states are shifted to lower energies and it decreases the permitted light absorption or emission frequencies. Thus, with the presence of an electric field, it is found that exciton still exists in the lower band gap material. This technique is particularly useful in Quantum confined stark effect (QCSE) optical modulators in which the optical communication signals are allowed to be switched ON and OFF quickly. Even if wells, dots or discs (quantum objects) emit or absorb light with higher energies than the material bandgap, the quantum confined stark effect will transfer the energy to the values which lower the bandgap. This was proved in the study of quantum discs which are embedded in a nanowire.

4.22 QUANTUM DOT LASERS

Quantum dot laser is a semiconductor laser that uses quantum dots in the active laser medium. Due to strong confinement of charge carriers in quantum dots, they provide an electronic structure similar to atoms. This increased carrier confinement would give different preferences to diode lasers over bulk materials.

The main features of quantum dot lasers are as follows:

- It offers low threshold current, better dependability with temperature, lower diffusion of carriers to the device surfaces, higher gain and a narrower emission line than two-fold hetero structure or quantum well lasers.
- (ii) It finds application in medicine (e.g. optical coherence tomography), display technologies (e.g. laser TV, projection), spectroscopy and optical data transmission system (e.g. optical LAN and metro-access system).
- (iii) It is capable of high speed operation at 1.3 μm wavelengths, at temperature from 20°C to 70°C i.e., greater stability with operation temperature.
- (iv) Its extensive gain spectrum makes them valuable for telecommunication amplifiers and tunable lasers.

In spite of the various points of interest expected for quantum dot lasers, their improvement has been thwarted by the challenges found in the manufacture of varieties of quantum dots free of imperfections and with uniform sizes. The best strategy to date has been simply the development of self-assembled quantum dots at the interface of two lattice mismatched materials.

Figure 4.33 shows a schematic diagram of an edge-emitting laser in view of self-assembled quantum dots. The device comprises of a few layers forming a pin diode structure. The layers are from bottom to top with an *N*-GaAs substrate, an *N*-AlGaAs layer, an intrinsic GaAs layer with InAs dots, a *P*-AlGaAs layer, and a *P*-GaAs top layer. Metallic contacts on the substrate and the top layer connect the device to an external circuit. Under the condition of forward bias voltage, electrons and holes are infused into the middle intrinsic GaAs layer or active layer, where they fall into the quantum dots, which have a small band gap and recombine there. The emission wavelength relates to the inter band transition of the InAs quantum dots.

The GaAs layer, which is sandwiched between AlGaAs layers with a lower refractive index, limits the light and expands the interaction with the carriers. The InAs wetting layer adds to an effective dispersion of carriers into the dots. Its bandgap is smaller than that of GaAs, and subsequently, gathers carrier that reach the GaAs layer. But its band gap is larger than that of the quantum dots since the wetting layer is very thin and carrier diffuse rapidly into the dots. In order to expand the density of quantum dots, a few wetting layers with pyramidal quantum dots are developed progressively over each other with a layer of GaAs in between to form a heap of quantum dots.



Fig. 4.33 *Quantum dot laser based on self-assembled dots with the insert showing wet layer with InAs pyramidal quantum dots*

4.23 OPTICAL DATA STORAGE TECHNIQUES

The technique used to write and read information in an external memory by light is called optical data storage technique. Lasers are used to write and read from small disks that contain a light-sensitive layer, upon which data can be stored. Storage can be achieved using lasers to pattern a surface, such as on a compact disk, or altering the physical properties of a small volume inside a light sensitive material. The advantage

of the optical data storage devices is that the disks are small and portable. They do not wear out easily with continuous usage.

Optical storage medium should have the ability to store the pattern of light intensity. Upon stimulation, the medium presents back a pattern of light intensity identical to the previously stored pattern. Consider the wavelength of light as λ , the area necessary to store a bit optically is about λ^2 . If A is the area of the disk in which the optical information is stored, then the maximum storage capacity SC_{max} is given by

$$SC_{max} = \frac{Area \text{ of the disk}}{\lambda^2}$$
 (4.27)

4.23.1 Disk Data Storage

Figure 4.34 shows the basic structure of recorded optical disk storage. Here, the pit is the smallest etched portion on the disk with very small volume. The recorded signal is encoded in the length of the pit and the spacing of pits is along the track. Here, the distance between two adjacent tracks i.e., track pitch is 1.6 μ m. The width of a pit is equal to a recording spot size of 0.5 to 0.7 μ m.

The light source used in an optical disk system is usually a GaAlAs semiconductor laser diode with wavelength from 0.78 to 0.83 μ m. The spot size of the read out beam is determined by the numerical aperture (N.A.) of the objective lens, by which the read out beam is focused on a pit. Typically,



 λ /N.A. = 1.55 is taken so that the effective diameter of the readout spot is approximately 1 λ m. A single readout spot does not cover two tracks or two pits since the spot size is larger than the width of a pit.

4.23.2 Recording and Readout of Information in CD-ROM

The reading and writing processes are shown in Fig 4.35. Using A/D converter, the information signal is first converted to binary coded signals. These pulses produce powerful laser pulses with high amplitude or intensity modulation. These emitted laser pulses produce pits on the rotating master CD. The depth, width and length of the pit produced by hole burning are determined by the intensity of the modulated laser pulse. The pits are elliptical in nature. The width of the laser spot is about 1 pm. The etched pit has a length of 930 nm, width of 500 nm and depth of 110 nm. Hence, the etched pits appear in the form of points. There are about 20,000 tracks over a 33 mm width on the CD. The etched points are covered by a thin layer of silver by electroplating. Then, this disk is pressed on the nickel disk. This nickel disk is converted into an electrically charged plate which is called master CD or mother CD. From mother CD, any number of CDs (sons) can be produced without change in fidelity or clarity. The master CD is pressed on a plastic disk on which the pits are produced in the same manner as in the master CD.

The optical disks are made using polymethyl methacryolite (polycarbonate plastics). Here, the recorded plastic disk is coated with aluminium to increase its reflectivity. Further, there is a protective polymeric coating to avoid formation of dusts and any penetration of moisture into the tracks. A circular hole is formed at the centre of the disk in order to insert it into the disk drive and it is placed into square type of plastic box. Thus, a recorded CD is produced.



Fig. 4.35 *Reading and writing in optical data storage device*

To get the playback information, the given CD is inserted into the CD player. A low power laser (GaAlAs laser) beam is incident on the pit of the CD and it is reflected back. The intensity of the reflected light is determined by the pit dimensions. When there is no pit, the incident light is fully reflected back to the photo detector. The reflected light and incident light will move in different directions without any overlap. The detector generates the electrical impulses based on the intensity of the reflected light. Further, these are converted into analog signal using a Digital to Analog (D/A) converter. The information can be recorded on one side only in a CD.

The recorded information in a CD can be obtained back by non-contact method. If there is no formation of scratches in CD during play back, then it leads to high life time. It has greater modulation bandwidth due to optical modulation occurring at very high frequencies. Hence, one can store a large amount of information in a CD.

4.23.3 Different Types of Optical Disks

There are different types of optical disks and they are discussed in detail in this section.

(a) Read only optical disk (CD-ROM)

For a read-only optical disk, such as compact disk (CD) for music recording or a compact disk read only memory (CD - ROM) for computers, the recorded data cannot be changed after the disk is manufactured. Since they store data optically, CD ROMs have a much higher memory capacity than computer hard disks that store data magnetically. However, CD ROM drives can only read information from the disk and cannot write on it.

The process of optical pick up is explained as follows: When there is no pit, the light is fully reflected to the detector. When there is a pit, the focussed beam covers both pit and the surrounding land. Both pit and land are coated with high reflective material like aluminium. So, the reflected light from a pit and the surrounding

land has a phase difference of π . Consequently, there is destructive interference at the detector and less light or no light is detected. The depth of the pit is typically 0.13 µm. Therefore, the recorded binary information is read optically with maximum light intensity and minimum light intensity.

Electronics and software interpret these data and access the information accurately from the CD ROM. The CD ROMs are popular for storing data buses and multimedia material. The most common format of CD-ROM holds data up to 600 MB while ordinary floppy disk holds only 1.44 MB data. CD - ROM players or drives spin the disk to access a sector of data and copy it into main memory for use by the computer. Further, CD-ROM players are available with different speed to read the data from the CD ROM. A single speed CD ROM player can read up to 1,50,000 bytes per second. Double speed, triple speed, quadruple speed, six-times speed and eight-times speed CD ROM players are currently available.

The important characteristics of CD-ROM players are their seek time and data transfer rate. The seek time (access time) measures how long it takes for the laser to access a particular segment of data. A typical CD-ROM player takes about a third of a second to access data as compared to a typical hard drive which takes about 10 ms to access data. The data transfer rate measures how quickly data is transferred from the disk to the computer's main memory.

(b) WORM optical disk

A WORM (write once, read many times) disk consists of either a poly-carbonate or hardened glass substrate and the recording layer consists of a highly reflective substance made of tellurium alloy. Laser beam is used to record data sequentially in WORM disk. A write beam burns a hole (pit) in the recording medium to produce a change in the reflectivity. In contrast to the CD ROM, the WORM optical disk directly modulates the reflected intensity of the readout beam. Hence, the WORMS are blank, writable compact disks.

(c) Magneto-optic (MO) disk

The MO disk combines both magnetic and optical data storage. Here, the users can record, erase and save data to these disks any number of times using special MO drives. It makes use of a MO recording material at room temperature and it is resistant to changes in magnetization. The magnetization of the material can be changed by placing it in a magnetic field and heating it to its Curie point about 180°C. To erase and to write new data on a MO disk, the heat of the write laser beam brings the recording material to its Curie point. Then, a bias magnet reverses the magnetization of the heated area that represents a bit. A low power linearly polarized laser beam can be used to read the data on the MO disk.

According to the Kerr magneto-optic effect, the polarization of the reflected readout beam will be rotated to the left or right depending on whether the magnetization of the recording material is upward (ones) or downward (zeros). The typical value of rotation of the plane of polarization is less than 1°. Manganese Bismuth (MnBi) medium in thin film form acts as a storage or memory element. The angle of rotation θ of the plane of polarization of light is directly proportional to the value of magnetization.

(d) Holographic optical data storage

Holography is the process of image construction by recording and reconstruction of hologram by means of interference technique. A hologram is a recording of the optical interference pattern that forms at the intersection of two coherent optical beams. Typically, light from a single laser is split into two paths, the *signal path* and the *reference path*. The beam that propagates along the signal path carries information, whereas the reference is designed to be simple to reproduce. A common reference beam is a *plane wave* which represents a light beam that propagates without converging or diverging. The two paths are overlapped

on the holographic medium and the interference pattern between the two beams is recorded. A key property of this *interferometric* recording is that when it is illuminated by a readout beam, the signal beam is reproduced. In effect, some of the light is diffracted from the readout beam to reconstruct a weak copy of the signal beam. If the signal beam was created by reflecting light off a 3D object, then the reconstructed hologram makes the 3D object appear behind the holographic medium.

(e) Digital Versatile Disk (DVD)

The CD technology has taken in a new direction with the advent of DVDs. DVD is a high density medium that is capable of storing a full length movie on a single disk of same size as that of a CD. It uses both sides of the disk. Each side can hold up to 4.7 GB. Like a CD, DVD has read only, recordable, and rewritable applications. The main feature of DVD-ROM is its enhanced multimedia and games applications.

(f) Blu-Ray Disk (BD)

Blu-Ray is a digital optical disk data storage format mainly designed to supersede the digital video disk (DVD) format. It is capable of storing hours of video in high-definition and ultra high-definition resolution. The main application of Blu-ray is as a medium for storing video content and also for the physical distribution of video games. The name *Blu-ray* refers to the blue laser used for reading the disk, which allows information to be stored at a greater density than the longer-wavelength red laser used for DVDs.

4.23.4 Advantages and disadvantage of Optical Storage Disk

The advantages of optical storage disk are as follows:

- (i) They are non-volatile which means that the memory can be retained even when the power is turned OFF.
- (ii) The speed of the optical data storage is high due to its random access ability.
- (iii) The durability is high. With proper care, optical media can last a long time, depending on choice of optical media.
- (iv) Several forms of optical media are write-once read-many and it is excellent for archiving because the data is preserved permanently with no possibility of being overwritten.
- (v) Transportability of the CD is easy. Optical media are widely used all platforms, including the computer.
- (vi) Optical media provide the capability to pinpoint a particular piece of data stored on it, independent of the other data on the volume or the order in which that data was stored on the volume.

The disadvantages of optical storage disk are as follows:

- (i) Optical disks require special drives to read/write.
- (ii) Optical storage does not provide enough data storage in comparison to other storage technologies
- (iii) Optical storage rewritable formats suffer from compatibility issues between drives.
- (iv) Optical storage is expensive per GB/TB in comparison to other technologies including disruptive formats that offer convenience over cost.
- (v) Consumer based formats like CD-R and DVD-R, BD-R do not have a formal method to grade the quality of disks since there are no standards for longevity tests.

TWO MARK QUESTIONS AND ANSWERS

1. Define electromagnetic radiation.

Electromagnetic radiation is a wave comprising of time varying electric E_x and magnetic field B_y components that are perpendicular to each other and also the direction of propagation is in z-direction.

2. Mention the frequency, wavelength and bandgap range of visible light.

Frequency range: 430 to 750 THz Wavelength range: 400 to 700 nm Bandgap range: 1.8 eV to 3.1 eV

3. Write the relation between electromagnetic constant and the electrical and magnetic constants. The electromagnetic constant and the electrical and magnetic constants are related by

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

where c is the velocity of light, μ_0 is the magnetic permeability and ε_0 is the electric permittivity in free space, respectively.

4. How frequency and wavelength are related to velocity?

The frequency v and the wavelength λ of the electromagnetic radiation are related to velocity by $c = \lambda v$.

5. Define refractive index.

The refractive index n of the medium is defined as the ratio of the velocity of light in free space c to the velocity in a medium v. It is given by

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

6. What are called photons?

In quantum-mechanical perspective, the electromagnetic radiation can be considered as groups or packets of energy, rather than as waves. These energy packets or particles are called *photons*. A photon is a massless and chargeless particle. The photon energy E is said to be quantized or can only have specific values as defined by

$$E = hv = \frac{hc}{\lambda}$$

where *h* is the Planck's constant, which has the value of 6.626×10^{-34} J·s.

7. What is meant by luminescence?

The general property of light emission is termed as luminescence. The phenomenon by which a material absorbs energy and then spontaneously emits visible or near-visible radiation is known as luminescence.

8. Define photoluminescence and electroluminescence.

When excess electrons and holes are created by photon absorption, photon emission from the recombination process is called photoluminescence.

Electroluminescence is the process of generating photon emission when the excitation of excess carriers is a result of an electric current caused by an applied electric field.

9. Give the classification of optical materials.

Based on the interaction with visible light, optical materials are classified into the following categories:

- (i) Transparent materials
- (ii) Translucent materials
- (iii) Opaque materials

10. What are called transparent, translucent and opaque materials?

Materials that transmit light with minimum absorption and reflection are known as *transparent materials* i.e., objects are easily seen through such transparent optical materials.

Materials in which light is transmitted diffusely are known as *translucent materials* i.e. objects are not clearly distinguishable when they are seen through such translucent optical materials.

Materials that do not allow the transmission of light through them are known as *opaque materials* i.e., these materials either absorb or reflect the entire light rays or photons and hence, become opaque.

11. What is the mechanism involved in carrier generation and recombination process?

Absorption and emission of light is one of the important mechanisms involved in carrier generation and recombination.

12. Define refraction.

When light is incident on a transparent material like a prism, the path of the light ray is deviated with a reduction in its velocity. This is known as *refraction*.

13. Write the relation between refractive index and dielectric constant of the material

The refractive index *n* is related to the dielectric constant ε_r of the material as $n \cong \sqrt{\varepsilon_r}$.

14. State Lambert's law.

Lambert's Law states that the absorbance of the material is directly proportional to its thickness (path length).

i.e., $I = I_0 e^{-\alpha x}$

where I_0 is the incident beam intensity, I is the transmitted beam intensity, x is the distance of light penetration into material from a surface and α is total linear absorption coefficient in m^{-1} . Here, α takes into account the loss of intensity due to scattering and absorption but it approaches zero for a pure insulator and hence, $I = I_0$.

15. What are called photodetectors?

Semiconductor devices that are used to detect the presence of photons and convert optical signals into electrical signals are called *photodetectors*. The *PN* junction forms the basis of photodetectors which include the photodiode and the phototransistor.

4.40 Physics for Electronics Engineering and Information Science

16. Give the performance parameters of photodetector.

The performance of photodetectors can be determined in terms of *figure of merit*. They are (i) responsivity and (ii) quantum efficiency.

17. List the types of photodetectors.

There are five main types of photodetectors. They are (i) Photoconductor, (ii) Photodiode, (iii) PIN Photodiode, (iv) Avalanche Photodiode and (v)Phototransistor.

18. Define photocurrent.

The current generated in a semiconductor device due to the flow of excess carriers generated by the absorption of photons is called photocurrent.

19. Write the principle of operation of photodiode.

A photodiode is a *PN* junction or PIN structure. When a photon of sufficient energy strikes the diode, it excites an electron, thereby creating a mobile electron and a positively charged electron hole. If the absorption occurs in the junction's depletion region, or one diffusion length away from it, these carriers are swept from the junction by the built-in field of the depletion region. Thus, holes move toward the anode, and electrons toward the cathode, and a photocurrent is produced.

20. Mention the applications of photodiode.

Photodiodes are used as light detectors, demodulators and encoders. They are also used in optical communication system, high-speed counting and switching circuits. Further, they are used in computer card punching and tapes, light operated switches, sound track films and electronic control circuits.

21. Give the applications of phototransistor.

The phototransistors find extensive applications in high-speed reading of computer punched cards and tapes, light detection systems, light operated switches, reading of film sound track, production line counting of objects which interrupt a light beam, etc

22. Write down the significance of Optocoupler.

Optocouplers are most often used to separate two circuit elements that are operating on extremely different voltages, which prevent damage to the part working at a lower voltage. They also work to keep the two elements from being damaged by reverse voltage or power surges. Because of these characteristics, optocouplers are best utilized in associated with ON/OFF switches and the transfer of digital data. They are commonly found between a transmitter and a receiver in an electric circuit.

23. What is photovoltaic effect?

If the *PN* junction is open circuited, the light energy is used to create a potential difference which is proportional to the frequency and intensity of the incident light. This phenomenon is called *photovoltaic effect*.

24. What is meant by photovoltaic cell?

A solar cell, or photovoltaic cell, is an electrical device that converts the energy of light directly into electricity by the photovoltaic effect, which is a physical and chemical phenomenon.

25. Under what principle does a photovoltaic cell work?

The photovoltaic effect is the creation of voltage or electric current in a material upon exposure to light and is a physical and chemical phenomenon.

26. Name two applications of photoconductive cells.

Photoconductive cells are used in different types of circuits and applications.

Analog Applications: Camera Exposure Control, Auto Slide Focus - dual cell, Photocopy Machines - density of toner, Colorimetric Test Equipment, Densitometer, Electronic Scales - dual cell, Automatic Gain Control - modulated light source, Automated Rear View Mirror

Digital Applications: Automatic Headlight Dimmer, Night Light Control, Oil Burner Flame Out, Street Light Control, Absence / Presence (beam breaker), Position Sensor

27. "A solar cell is a *PN* junction device with no voltage directly applied across the junction". If it is so, how does a solar cell deliver power to a load?

Solar cell consists of a single semiconductor crystal which has been doped with both *P*- and *N*-type impurities, thereby forming a *PN* junction. The incident sunlight passes through glass plate and reaches the *PN* junction. An incident light photon at the junction may collide with the valence electron and impart sufficient energy to make a transition to the conduction band. As a result, a electron-hole pair is formed. The newly formed electrons are minority carriers in the *P*-region. They move freely across the junction. Similarly, holes formed in the *N*-region cross the junction in the opposite direction. The flow of these electrons and holes across the junction is in a direction opposite to the conventional forward current in a *PN* junction. Further, it leads to the accumulation of a majority carrier on both sides of the junction. This gives rise to a photovoltaic voltage across the junction in the open circuit condition. In bright sunlight, about 0.6 V is developed by a single solar cell. The amount of power, the cell can deliver depends on the extent of its active surface. An average cell will produce about 30 mW per square inch of surface, operating in a load of 4 Ω .

28. Expand LASER.

LASER: Light Amplification by Stimulated Emission of Radiation

29. What is meant by population inversion and stimulated emission?

Population inversion is a non-equilibrium condition in which the concentration of electrons in one energy state is greater than that in a lower energy state.

The process by which an electron is induced by an incident photon to make a transition to a lower energy state, emitting a second photon is called stimulated emission.

30. What is called radiative recombination?

The recombination process of electrons and holes that produces a photon, such as the direct band-toband transition in gallium arsenide is called radiative recombination.

31. List out the materials responsible for the colour of emitted light in LED.

The colour of the emitted light depends on the type of material used is given as follows:

- (iv) Gallium arsenide (GaAs) infrared radiation (invisible)
- (v) Gallium phosphide (GaP) red or green
- (vi) Gallium arsenide phosphide (GaAsP) red or yellow

4.42 Physics for Electronics Engineering and Information Science

32. Mention the applications of LED.

LEDs are used in burglar alarm systems, picture phones, multimeters, calculators, digital meters, microprocessors, digital computers, electronic telephone exchange, intercoms, electronic panels, digital watches, solid state video displays, and optical communication systems.

33. Write the two types of OLED.

OLED are of two types. They are (i) Passive matrix organic LED (PMOLED) display and (ii) Active matrix organic LED (AMOLED) display.

34. What is meant by TFT backplane?

A vital technology which majorly sets the limits of display industry is known as the *backplane technology*. The backplane is responsible for turning ON and OFF of each pixel. It is constructed using thin film transistor (TFT) that acts as a switch. So, it is generally called TFT backplane. The power consumption of TFT backplane is very low. This gets refreshed fast in OLED and there is increased convenience for TV and other displays which broadcast moving pictures and graphics.

35. What are excitons?

Excitons are electron-hole pairs that bound together in stable orbits by the Coulomb force of attraction between them. They are usually found in materials like insulators, semiconductors and some liquids that have low dielectric value.

36. Mention the types of excitons.

Excitons are of two types. They are (i) Frenkel exciton and (ii) Wannier exciton.

37. Define Frenkel exciton.

Excitons that are tightly bounded and localized on individual atom sites are known as Frenkel excitons. These excitons have very small radii and large binding energies.

38. Define Wannier exciton.

Excitons that are weakly bounded and move freely throughout the ionic crystal are known as Wannier excitons. These excitons have large radii and possess a weak electron-hole interaction due to valence electrons.

39. What is meant by quantum confined stark effect?

Quantum confined stark effect is the shift in the quantum levels induced by an electric field that causes a shift in the band edge and exciton energies. This effect is mainly used in optical modulators.

40. What is quantum dot laser?

Quantum dot laser is a semiconductor laser that uses quantum dots in the active laser medium. Due to strong confinement of charge carriers in quantum dots, they provide an electronic structure similar to atoms. This increased carrier confinement would give different preferences to diode lasers over bulk materials.

41. Mention the different types of optical disks.

There are different types of optical disks. They are (i) Read only optical disk(CD - ROM), (ii) WORM optical disk, (iii) Magneto-optic (MO) disk, (iv) Holographic optical data storage, (v) DVDs and (vi) Blu-Ray disks

REVIEW QUESTIONS

- 1. Write the frequency, wavelength and bandgap energy ranges of visible light.
- 2. Give the graphical representation of electromagnetic spectrum.
- 3. Draw the spectral response of human eye and explain briefly.
- 4. What is the distinction between *fluorescence* and *phosphorescence*?
- 5. Explain the classification of optical materials.
- 6. Describe the carrier generation and recombination process in semiconductors.
- 7. Briefly explain why metals are opaque to electromagnetic radiation having photon energies within the visible region of the spectrum.
- 8. Explain the optical properties of metals.
- 9. Describe the optical properties in insulators.
- 10. Briefly explain what determines the characteristic colour of a metal and a transparent insulator.
- 11. Discuss the emission and absorption of light in semiconductors.
- 12. Explain the principle and working of photodiode.
- 13. Explain the working principle of photodetectors.
- 14. Discuss the figure of merit of photodetectors.
- 15. Define photoconductivity.
- 16. Write the current equation for the volt-ampere characteristics of a photodiode. Define each term in the equation.
- 17. Explain the volt-ampere characteristics of a semiconductor photodiode.
- 18. List the applications of a photodiode.
- 19. Write notes on (i) *PIN* photodiode and (ii) Avalanche photodiode.
- 20. Discuss the advantages of PIN and avalanche photodiodes compared to the simple PN junction photodiode.
- 21. Describe with neat diagrams, the operation of a phototransistor and state its applications.
- 22. With output characteristics, explain how phototransistor responds to the incident light.
- 23. Explain the difference between a photodiode and photovoltaic cell.
- 24. Give the constructional features of a solar cell and explain its principle of operation. Mention two of its applications.
- 25. Distinguish between photovoltaic cell and solar cell.
- 26. State the applications of photovoltaic cell.
- 27. Explain how solar cells can be used for the conversion of solar energy directly to electricity.
- 28. List the applications of optocoupler.
- 29. Describe with the help of a relevant diagram, the construction of an LED and explain its working.
- 30. List the applications of an LED.
- 31. Compare the working principle of LED with solar cell.
- 32. In what respect is an LED different from an ordinary PN junction diode?
- 33. Describe the working of OLED.
- 34. Mention the applications of AMOLED and PMOLED.
- 35. Explain the principle behind the laser diode with a neat sketch.
- 36. Discuss the concept of population inversion in a laser diode.
- 37. Compare the characteristics of LED and LASER diode.
- 38. Explain the properties of excitons.

4.44 Physics for Electronics Engineering and Information Science

- 39. Describe quantum confined stark effect.
- 40. Discuss the principle of quantum dot lasers.
- 41. Write notes on the optical data storage technique.
- 42. Discuss about the different types of optical storage disks.

5

Nanoelectronic Devices

5.1 INTRODUCTION

Electronic devices with at least one dimension in nanometer scale (≤ 100 nm) are called nanoelectronic devices. Scaling down of sizes has enormous impact on all the properties of the device. As sizes are scaled down, length and surface related phenomena become more significant as $L > L^2 > L^3$ for small values of L. Therefore, phenomena like diffusion, heat dissipation, conductivity, friction and viscosity become dominant in nanodevices. When charge carriers are confined to small structural dimensions comparable to their de Broglie wavelengths, they exhibit different electrical, chemical, magnetic, optical and mechanical properties when compared to their macroscopic counterparts. While macroscopic electronic devices deal with manipulating motion of group of electrons, nanodevices are heading towards controlling the motion of a single electron. Therefore, nanoelectronic devices open up possibility of new and interesting applications in a wide variety of fields.

Examples of unwitting use of nanotechnology can be traced to ancient times, in the fourth century stained glass paintings in medieval Roman churches, the wootz steel blades used nearly 2000 years ago in India to make the famous Indian Damascus swords and the old carbon-rich paintings of Ajanta caves in India. However, the first scientific discussion on the possibility of implementation of nanostructures and their far reaching implications was by Feynman in his famous talk on *"There is plenty of room at the bottom"* at the annual meet of American Physical Society in 1959. Ever since then, advances in this field have occurred in a rapid pace.

Nanostructures have been realised either through (i) a top-down approach, where the bulk material has been selectively etched to realise nanodevices using advanced lithography techniques or (ii) using the bottom-up approach where nanostructures are grown through layer–by-layer molecular or atomic assembly. Microelectronic industry has adopted the top-down approach to realise commercial nanoelectronic devices. Some of these have already become a part of our day-to-day life, in the form of ultra-thin, light weight mobiles, high speed computers, point-of-care bio-devices and sensors for Internet of Things (IoT).

In nanoelectronic devices, charge carriers, i.e., electrons are either confined or interact with structures with sizes comparable to their matter waves. Hence, their physical properties need to be modelled using quantum mechanics. At these dimensions, the effects of discrete nature of electron transport, low scattering in the solids and spin state of electron start affecting the electron transport characteristics. Therefore, this chapter

discusses some of the transport phenomena and the basic blocks of nanoelectronic devices like quantum wells, quantum wires and quantum dots.

5.2 ELECTRON DENSITY IN BULK MATERIALS

The density of electrons n_e available for conduction in a bulk material at any temperature T is given by the product of the density of states, N(E) available in the material, effective mass of an electron m_e^* and the probability f(E), that they are occupied by electrons at that temperature, integrated over the energy range of interest.

$$n_e = \int_E N(E) f(E) dE$$
(5.1)

where $N(E) = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \sqrt{E}$ and $f(E) = \frac{1}{1 + e^{\left(\frac{E - E_F}{kT}\right)}}$

This quantum mechanical expression holds good for all materials. At the temperature 0 K, the highest occupied electronic state is the Fermi level $E_{F_{,}}$ and the energy of this state is defined as the Fermi energy. Hence, the electron density n_e at 0 K is obtained by integrating Eqn. (5.1) between the limits 0 to E_F and is given by

$$n_e = \frac{2}{3}\gamma E_F^{3/2} = \frac{8\pi}{3h^3} (2m_e^* E_F)^{3/2}$$
(5.2)

Therefore, the number of electrons, N becomes $\frac{8\pi}{3h^3}(2m_e^*E_F)^{3/2}L^3$, where L represents dimension of the

solid. Thus, it is seen that the number of electrons decreases drastically as dimensions scale down.

5.3 SIZE DEPENDENCE OF FERMI ENERGY

The expression for Fermi energy can be obtained by using Eqn. (5.2). It is obvious that the dependence of this energy on structural dimension is only through the density of electrons, $n_e = N / L^3$.

$$E_F = \left(\frac{3n_e}{2\gamma}\right)^{2/3}$$

$$E_F \propto \frac{1}{I^2}$$
(5.3)

Therefore,

So, the Fermi energy increases as size of the material decreases. The Fermi energy level, at any temperature higher than 0 K, will lie at the bottom of the conduction band in a metal and will move into the conduction band as temperature increases whereas in semiconductors and insulators, the Fermi level generally lies in the bandgap region.

Referring to Eqn. (1.70) in Chapter 1, the energy of each state for electron E_n confined in 3D is given by

$$E_n = \frac{n^2 h^2}{8m_e^* L^2}$$
(5.4)

Both E_n and the energy separation between the states ΔE , scale as $1/L^2$. Thus, as sizes scale down, the energy of each state and separation between the energy levels increase and number of electrons decreases.



Fig. 5.1 Schematic of change in energy levels with different sizes of the material

The quasi-continuous density of states in the valence and the conduction bands of bulk material will break into discrete electronic levels as size reduces. The spacing between these levels and the band gap increases with decreasing particle size as shown in Fig. 5.1.

The energy bands of semiconductor gradually change into discrete molecular electronic levels. The charge carriers may be treated quantum mechanically as "particles in a box" when the particle size is less than the de Broglie wavelength of the electrons. Here, the size of the box is given by the dimensions of the crystallites. The bandgap in semiconductors is found to increase for clusters ranging from 1 nm to almost 10 nm. The metal particles with 50 to 100 atoms having a diameter between 1 and 2 nm begin to loose their metallic behaviour and tend to become semiconductors. Such particles that show this size based quantization of energy states are referred as Q-particles or quantum dots and it will be discussed later in this chapter. These small particle sizes are also called clusters.

5.4 BANDGAP OF NANOMATERIALS

A nanomaterial has only very few atoms. As the size of the nanomaterials is very small, important phenomena such as quantum-confinement effects, insufficient material effects, dominance of surface plasmons are present. As shown in Fig. 5.2, the energy states are discrete in atoms. When two atoms come closer, the sub levels of the energy state split and broaden into bands, as more atoms are added. As bands get broader in solids, the band gaps between them become smaller. Nanoparticles represent a state of matter in the transition region between bulk solid and single molecule. Their physical and chemical properties gradually change from solid state to molecular behaviour with decreasing particle size. Due to their small dimensions, the surface-to-volume ratio increases, and the number of surface atoms may be similar to or higher than those located in the crystalline lattice core and the surface properties are no longer negligible. When no other molecules are absorbed onto the nano-crystallites, the surface atoms are highly unsaturated and their electronic contribution to the behaviour of the particles is totally different from that of the inner atoms. This leads to different electronic transport and catalytic properties of the nanoparticles.

The second phenomenon, which occurs in metal and semiconductor nanoparticles, is totally an electronic effect. The band structure gradually evolves with increasing particle size, i.e., molecular orbital gets converted into delocalised band states. Figure 5.2 shows the size quantization effect responsible for the transition between a bulk solid and nanoparticle. In a metal, the quasi-continuous density of states in the valence and

the conduction bands splits into discrete electronic levels. The spacing between these levels and the bandgap increases with decreasing particle size. When the particle size is less than de Broglie wavelength of electrons, the charge carriers can be treated as particles in a box.



Fig. 5.2 Size quantization effect on the band gap

In nanoparticles, as the size is decreased, electrons get confined to the particles. This leads to the following changes in the properties of nanoparticle.

- (i) Bandgap energy is increased
- (ii) Band levels get discretized.

The effective bandgap energy of semiconductor nanoparticle of radius R is given by

$$E_{G}^{\text{effective}}(R) = E_{G}(\infty) + \frac{h^{2}\pi^{2}}{2R^{2}} \left(\frac{1}{m_{e}^{*}} + \frac{1}{m_{h}^{*}}\right) - \frac{1.8e^{2}}{\varepsilon R}$$
(5.5)

where $E_G(\infty)$ is the bandgap of bulk solid, ε is dielectric constant of bulk solid, m_e^* is the effective mass of electron and m_h^* is the effective mass of hole. From the above equation, it is seen that as the radius decreases, the effective bandgap of nanoparticle increases.

5.5 QUANTUM CONFINEMENT

In general, if the volume of the material is reduced such that energy levels appear as discrete levels within each band, the number of electrons or charges confined in these volumes decreases and becomes quantised and countable. This confinement, which results in few countable numbers of electrons or charges, is therefore called *quantum confinement*. Quantum confinement can be realised within a bulk material by creating isolated regions of small volume or localised regions. The methodology for realising this practically, is to create small volumes of semiconductors or conductors enclosed by insulators, or by confining the electrons in a small region through applied electric fields. Microelectronic industry exploits this top-down approach to fabricate quantum confined structures for nanoelectronic applications.

When a thin layer of material with low bandgap (e.g., GaAs) is sandwiched between materials of comparatively higher bandgap (GaAlAs) to form a heterojunction, the carriers in the low bandgap (GaAs) layer are subjected to a potential barrier at the boundary as shown in Fig. 5.3. This confines their motion within its thin layer creating a quantum confined structure.



Fig. 5.3 Illustration of band diagram of a double heterojunction

The amount of size reduction required to achieve quantum confinement can be calculated from knowledge of the lowest allowed energy level of an electron in a 3D potential well. The lowest allowed energy level of an electron trapped in a 3D potential well is

$$E_{111} = \frac{h^2}{8m_e^*L^2}$$

Therefore, if the size of the structure is reduced, such that, only the lowest energy E_{111} is possible, then this provides the volume at which quantum confinement of a single electron can be realised. Thus, single electron devices can be realised when material volume is scaled down to a cube of side L as given by

$$L = \sqrt{\frac{h^2}{8E_{111}m_e^*}}$$
(5.6)

The dimension L can also be expressed in terms of the de Broglie wavelength, λ_{dB} , of the electron in the lowest energy state to give $L = \lambda_{dB}/2$. In practise, quantum confinement is observed even for confined dimensions as large as a few multiples of λ_{dB} .

EXAMPLE 5.1

Calculate the critical dimension required to confine an electron in copper with ground state energy of 8eV.

Solution

Method 1: Kinetic Energy of the electron,
$$E = \frac{p^2}{2m_e^*} = 8 \text{ eV}$$

Therefore, momentum p of the electron with energy of 8 eV is

$$p = \sqrt{2m_e^*E} = \sqrt{2 \times 9.11 \times 10^{-31} \times 8 \times 1.602 \times 10^{-19}} = 15.27 \times 10^{-25} \text{ kg} \cdot \text{m/s}$$

De-Broglie wavelength of electron with energy of 8 eV,

$$\lambda_{\rm dB} = \frac{h}{p} = \frac{6.62 \times 10^{-34}}{15.27 \times 10^{-25}} = 4.33 \times 10^{-10} \,\mathrm{m}$$

Critical dimension required to confine this electron is

$$\frac{\lambda_{\rm dB}}{2} = \frac{4.33 \times 10^{-10}}{2} = 2.17 \times 10^{-10} \,\mathrm{m}$$

Method 2: Critical dimension =
$$\sqrt{\frac{h^2}{8E_{111}m_e^*}} = \sqrt{\frac{(6.62 \times 10^{-34})^2}{8 \times 8 \times 1.602 \times 10^{-19} \times 9.11 \times 10^{-31}}} = 2.17 \times 10^{-10} \,\mathrm{m}$$

5.6 QUANTUM STRUCTURES

To realise quantum confinement, bulk volume can be scaled down by reducing any of the three dimensions, or by reducing all three dimensions simultaneously. A cubical quantum structure can be realised by isolating one small cubical region of side tens of nm from the rest of the bulk material. One way to accomplish this is to etch off material around it leaving air as insulator. The other method is to deposit material of larger bandgap all round this region to restrict motion of the carriers. If the isolated region has two dimensions reduced to nano sizes, then it is called a nano wire, whereas if only one dimension is reduced, then it is called a quantum well. The different quantum structures made through top-down approach are shown in Fig. 5.4 along with those realised through bottom-up approach through growth.

Size based classification of nano materials is done depending on the number of dimensions that are in macroscopic scale. Thus, 2D materials are those in which two dimensions are in macroscopic scale and one dimension alone is < 100 nm. Examples of 2D material are thin films, nanowires, carbon nanotubes etc. 0D nanomaterials are those in which there are no dimensions that are macroscopic, i.e all three dimensions are < 100 nm. Fullerenes, quantum dots etc are some examples. Quantum wires and quantum wells allow electrons to move freely in one and two directions, respectively and therefore, they can be said to offer partial confinement of electrons whereas quantum dots offer full confinement. Metal quantum dots require critical dimensions of ~10 nm. Semiconductor quantum dots are easier to realise practically. This difference in the confinement occurs because the free electrons in metals need to be confined, whereas the excitons in semiconductors need to be confined. Exciton is a quasi-particle, which is a weakly bound state of electron and hole, created through columbic attraction, over a characteristic distance called exciton Bohr radius. The name Bohr radius has been given, comparing the exciton binding distance, to the radius of the electron orbital in a hydrogen atom. The quantum confinement of excitons requires that the dimension is reduced to one exciton Bohr radius, which is in the range of 3-20 nm.



5.7 DENSITY OF STATES IN QUANTUM STRUCTURES

Understanding the change in the density of states in quantum structures is essential to calculate the charge densities and to predict the conduction phenomenon in them. To determine density of states, the number of allowed states per unit volume in an energy interval dE needs to be estimated, and then expressed in terms of energy.

5.7.1 Density of States in 2D Nanomaterial (Quantum Wells)

A thin layer of low bandgap material (GaAs) shown as shaded region is sandwiched between larger bandgap material (GaAlAs) as shown in Fig. 5.5. This will confine the motion of an electron to the GaAs layer along the x-direction. This is the typical configuration used in heterojunction lasers and is an example of 2D nanomaterial. The GaAs layer is a quantum well providing quantum confined motion of electron in x-direction. The electrons or charge carriers are free to move in the y and z-directions. The wavefunction in these directions



Fig. 5.5 2D quantum well

0

Fig. 5.6 2D quantum well

V = ∞

L

can be represented by plane waves and the motion of the charge carrier will be similar to that in a periodic potential. Their energy also becomes quantized. The wave function in the quantum confined *x*-direction will be the solution of the 1D Schrödinger equation in the quantum confined direction in the *x*-axis. Therefore, the number density of states in 2D nanomaterial per unit energy interval requires the calculation of number of allowed quantum confined states in *x*-direction and those allowed by the Bloch solutions in the other two directions.

Hence, the 3D wave function of the quantum well structure as shown in Fig. 5.6 can be written as

$$\Psi(x, y, z) = \Psi(x)e^{ik_y y}e^{ik_z z}$$
(5.7) V

For a particle confined in a one dimensional box, the possible solution for wave function, on applying boundary conditions that it should vanish at both x = 0 and $x = L_x$, is

$$\psi(x) = A\sin k_x x \tag{5.8}$$

where $k_x = \frac{n_x \pi}{L_x}$ with $n_x = 1, 2, 3...$ is quantized.

Bloch solutions in the y and z-directions result in quantization of k_y and k_z . They can take values given by $n_y 2\pi$ $n_z 2\pi$

$$k_y = \pm \frac{n_y 2\pi}{L_y}$$
 and $k_z = \pm \frac{n_z 2\pi}{L_z}$ where $n_y, n_z = 0, 1, 2...$

Total energy of the charge carrier, $E = \frac{\hbar^2 k_x^2}{2m_e^*} + \frac{\hbar^2 k_y^2}{2m_e^*} + \frac{\hbar^2 k_z^2}{2m_e^*} + \frac{\hbar^2 k_z^2}{2m_e^*}$ (5.9) Substituting k_x in Eqn. (5.8), we get

$$E = \frac{\hbar^2 n_x^2 \pi^2}{2m_e^* L_x^2} + \frac{\hbar^2 k_y^2}{2m_e^*} + \frac{\hbar^2 k_z^2}{2m_e^*} = \frac{\hbar^2 n_x^2 \pi^2}{2m_e^* L_x^2} + \frac{\hbar^2 (k_y^2 + k_z^2)}{2m_e^*}$$
$$= \frac{\hbar^2 n_x^2 \pi^2}{2m_e^* L_x^2} + \frac{\hbar^2 (k^2)}{2m_e^*} \qquad (\text{since } k^2 = k_y^2 + k_z^2)$$
$$E = E_{n_x} + E_{n_y n_z} \qquad (5.10)$$

or

 $E_{n_x n_z}$ can be represented in k-space by a circle as shown in Fig. 5.7.

Area occupied by a state in this k-space $=\frac{2\pi}{L_x}\cdot\frac{2\pi}{L_y}=\frac{4\pi^2}{A}$

Total area in k space corresponding to energy $E_{n_v n_z} = \pi k^2$

Therefore, number of states with energy $E_{n_y n_z} = \frac{\pi k^2}{4\pi^2 / A}$

Therefore, the number of states in y-z plane, n, including the 2 spin Fig. 5.7 states becomes

$$n = 2 \times \frac{k^2}{4\pi} = \frac{k^2}{2\pi}$$
(5.11)

The energy $E_{n_y n_z}$ associated with these states can be written as

$$E_{n_y n_z} = \frac{\hbar^2 (k^2)}{2m_e^*} = (E - E_{n_x})$$
(5.12)

The density of states i.e., number density of states per unit energy dn/dE can be written by the chain rule,

$$\frac{dn}{dE} = \frac{dn}{dk}\frac{dk}{dE}$$
(5.13)

From Eqn. (5.11),

$$dn = \frac{2kdk}{2\pi} = \frac{kdk}{\pi}$$
$$\frac{dn}{dk} = \frac{k}{\pi}$$
(5.14)

From Eqn. (5.12),

$$dE = \frac{\hbar^2 (2k)}{2m_e^*} dk = \frac{\hbar^2 k}{m_e^*} dk$$
$$\frac{dk}{dE} = \frac{m_e^*}{\hbar^2 k}$$
(5.15)

i.e.,

i.e.,

Substituting Eqns. (5.14) and (5.15) in Eqn. (5.13), the density of states in y-z plane for a particular value of E_{n_x} is given by

$$\frac{dn}{dE} = \frac{k}{\pi} \times \frac{m_e^*}{\hbar^2 k} = \frac{m_e^*}{\pi \hbar^2}$$
(5.16)

i.e., density of states increment by a quantity $\left(\frac{m_e^*}{\pi\hbar^2}\right)$ for each change in E_{n_x} .

Therefore, the general expression for density of states for the quantum well can be written as

$$N(E)dE = \frac{m_e^2}{\pi \hbar^2} \sum_{n_x} u(E - E_{n_x})$$
(5.17)



y and z-direction

where *u* represents a unit step function which is 1 when the argument is positive and 0 if argument is negative. Figure 5.8 represents the density of states for the quantum well. Even though the density in two dimensions is constant, the density of states for a quantum well is a step function with steps occurring at the energy of each allowed level of the quantum confined dimension.



Fig. 5.8 (a) Variation of density of states for a 2D nanomaterial (quantum well) with energy (b) energy levels of the quantum confined levels, momentum space representation of energy levels of the quantum well and the density of states.

5.7.2 Density of States in a 1D system (Quantum Wires, Nanowires)

The schematic of one dimensional quantum wire is shown in Fig. 5.9. The estimation of the density of states in a 1D system requires solving the Schrödinger equations to obtain the matter waves of the carrier, under quantum confinement in two dimensions, say x and y space, and the wave function in the z direction where the carrier is free to move. In z direction, the wave function can be represented by a plane wave and in this dimension, charge carrier is moving in a periodic potential and hence, Bloch solutions apply.



Fig. 5.9 One dimensional quantum wire

The complete representation of matter wave can be written as,

$$\psi(x, y, z) = \psi(x, y)e^{ik_{z}z}$$
(5.18)

The Schrödinger equation for the particle inside a 2D potential well of size $L_x = L_y = L$ where potential energy V = 0 can be written as

$$\frac{-\hbar^2}{2m_e^*} \left(\frac{\partial^2 \psi(x,y)}{\partial x^2} + \frac{\partial^2 \psi(x,y)}{\partial y^2} \right) = E\psi(x,y)$$
(5.19)

Using the method of separation of variables, wave function is expressed as

$$\psi(x, y) = \psi(x)\psi(y) \tag{5.20}$$

Substituting Eqn. (5.20) in Eqn. (5.19) and dividing throughout by $\psi(x)\psi(y)$, we get

$$\frac{1}{\psi(x)}\frac{\partial^2\psi(x)}{\partial x^2} + \frac{1}{\psi(y)}\frac{\partial^2\psi(y)}{\partial y^2} + k^2 = 0$$
(5.21)

where

$$k^2 = \frac{2m_e^*E}{\hbar^2}$$

The 1st and 2nd term of Eqn. (5.21) are the function of x and y respectively. This equation will be valid for all values of x and y only if $\psi(x)$ and $\psi(y)$ are independently equal to a constant. Therefore, expressing $k^2 = k_x^2 + k_y^2$, Eqn. (5.21) can be split into two independent equations as

$$\frac{1}{\psi(x)}\frac{\partial^2\psi(x)}{\partial x^2} + k_x^2 = 0 \quad \text{and} \quad \frac{1}{\psi(y)}\frac{\partial^2\psi(y)}{\partial y^2} + k_y^2 = 0$$

Solutions for the wave-functions $\Psi(x)$ and $\Psi(y)$ will be a combination of sine and cosine functions of k_x and k_y respectively. Therefore,

$$\psi(x) = A_1 \sin(k_x x) + B_1 \cos(k_x x)$$
 and $\psi(y) = A_2 \sin(k_y y) + B_2 \cos(k_y y)$ (5.22)

Applying boundary conditions at x = 0, x = L and y = 0 and y = L where the wave functions $\Psi(x)$ and $\Psi(y)$ should go to zero, rules out cosine solutions and the only possible values for *k* become,

$$k_x = \frac{n_x \pi}{L}, \ k_y = \frac{n_y \pi}{L}$$

where n_x and $n_y = 1, 2, 3...$

Thus, energy E in this system can be written as

$$E = E_{n_x n_y} + E_z = \frac{\hbar^2}{2m_e^*} (k_x^2 + k_y^2) + \frac{\hbar^2 k_z^2}{2m_e^*}$$
(5.23)

Substituting the values of k_x and k_y in Eqn. (5.23), we get

$$E = \frac{\hbar^2 \pi^2}{2m_e^* L^2} (n_x^2 + n_y^2) + \frac{\hbar^2 k_z^2}{2m_e^*}$$
(5.24)

The electron is free in the z-direction but its energy due to motion in the x-y plane is quantized and can take only discrete set of values. It is easy to count the number of states in the x and y directions.

$$E_{z} = \frac{\hbar^{2}k_{z}^{2}}{2m_{e}^{*}} = (E - E_{n_{x}n_{y}})$$
(5.25)

The space occupied by a state in this k-space (due to confinement in z-direction) is $\frac{2\pi}{L_z}$

The number of states $n(E_z)$ including spin $= 2 \times \frac{k_z}{2\pi / L_z} = \frac{L_z k_z}{\pi}$

Here,
$$\frac{dn}{dk} = \frac{L_z}{\pi}$$
 and from Eqn. (5.25), we have $\frac{dk}{dE} = \frac{2m_e^*}{\hbar^2} \frac{1}{2k_z}$

Therefore, density of states is

$$\frac{dn}{dE} = \frac{dn}{dk} \cdot \frac{dk}{dE} \cdot \frac{1}{L_z} = \frac{L_z}{\pi} \cdot \frac{2m_e^2}{\hbar^2} \frac{1}{2k_z} \frac{1}{L_z} = \frac{m_e^2}{\pi\hbar^2} \frac{1}{k_z} = \frac{m_e^2}{\pi\hbar^2} \frac{1}{\sqrt{2m_e^*(E - E_{n_x n_{xy}})}} \\
= \frac{1}{\pi\hbar} \sqrt{\frac{m_e^*}{2}} \sum_{n_x \cdot n_y} \frac{u(E - E_{n_x, n_{xy}})}{\sqrt{(E - E_{n_x, n_{xy}})}}$$
(5.26)

The variation of density of states for quantum wire with energy is shown in Fig. 5.10(a). Here, again the summation over n_x, n_y gives the incremental change that occurs at energies corresponding to each allowed

state of the two dimensional quantum confinement. The density of states for the quantum wire is complicated by the degeneracy of the energy levels: for instance a two-fold degeneracy increases the density of states associated with that energy level by a factor of two as shown in Fig. 5.10(b).



Fig. 5.10 (a) Variation of density of states for a 1D nanomaterial (quantum wire) with energy and (b) energy levels of the quantum confined levels, momentum space representation of energy levels of the quantum well and the density of states

The dashed curve in Fig. 5.10 (a) shows a gradual increase in the minimum density of states which is mainly due to the step-like increase in the density of states at each allowed energy level of the 2 quantum confinement dimensions. This can be worked out by looking at the density of states in the quantum confined x-y plane.

Number density of states in the x-y plane

The equation for energy $E_{n_x n_y}$ in x-y plane represents a circle as shown in Fig. 5.11 in the momentum space (taking the first term of 5.24). The unit area of this space contains one allowed state.

Therefore, area occupied by one state $=\frac{\pi}{L} \times \frac{\pi}{L} = \frac{\pi^2}{L^2}$

The total area in k space corresponding to energy $E_{n_x n_y}$ is πk^2 and therefore number of allowed states for energy $E_{n_x n_y}$ is

therefore, number of allowed states for energy $E_{n_x n_y}$ is



Fig. 5.11 Direction of quantization at x-y plane

$$n(E_{n_x n_y}) = \frac{\pi k^2}{\pi^2 / L^2} \times \left(\frac{1}{4}\right) \times 2 = \frac{k^2 L^2}{2\pi}$$
(5.27)

where (1/4) accounts for taking only the positive quadrant as *n* takes only positive values and the factor of 2 takes into account the two electrons of opposite spins that can occupy one state.

Substituting for k^2 in the above equation, we get

$$n(E_{n_x n_y}) = \left(\frac{2m_e^* E}{\hbar^2}\right) \frac{L^2}{2\pi} = \frac{m_e^* E}{\hbar^2} \frac{L^2}{\pi}$$
 5.28)

The number of states in an interval dE is given by

$$\frac{dn}{dE} = \frac{dn}{dk} \cdot \frac{dk}{dE} = \left(\frac{2kL^2}{2\pi}\right) \times \left(\frac{m_e^*}{\hbar^2 k}\right) = \frac{L^2 m_e^*}{\pi \hbar^2}$$
(5.29)

The density of states in the interval *dE* in the *x-y* plane is $\frac{L^2 m_e^* / \pi \hbar^2}{L^2} = \frac{m_e^*}{\pi \hbar^2}$

This is independent of E_{n,n_y} . Thus, the density of states in the quantum confined x-y plane increases by fixed amount.

Density of States in a 0D System (Quantum Dots) 5.7.3

The schematic of 0D system like quantum dots is shown in Fig. 5.12. In a 0D material, charged particles are quantum confined in all three directions. Therefore, these systems will have a series of discrete allowed states which will be represented by delta functions.



Fig. 5.12 Zero dimensional quantum dots

i.e.,
$$N(E)dE = \sum_{n_x, n_y, n_z} \delta(E - E_{n_x, n_y, n_z})$$



Figure 5.13 represents the density of states for the quantum dot.



(a) Variation of density of states for a OD nanomaterial (quantum dot) with energy (b) energy Fig. 5.13 levels of the quantum confined levels, momentum space representation of energy levels of the quantum well and the density of states

Figure 5.14 shows the density of states for bulk, quantum well, quantum wire and quantum dot structures.



Normalized energy densities for bulk 3D structures, quantum well (2D) structures, Fig. 5.14 quantum wires (1D) and quantum dots (0D) structures

5.8 ZENER-BLOCH OSCILLATIONS

A quantum confined electron, which moves without scattering, in a nanostructure with nearly perfect crystal structure at low temperatures, can exhibit different types of transport like (i) ballistic transport that follows Newton's laws for undamped systems (ii) Zener-Bloch oscillations while moving in a periodic potential under a constant DC external field and (iii) resonant tunneling through thin barriers under the influence of a bias voltage. Bloch has shown that when an electron moves through a periodic lattice subjected to an external field E, its propagation vector k changes with respect to time as per the quantum mechanical representation of Newton's law.

i.e., $\hbar \frac{dk}{dt} = eE$ (5.31) The electron starts from the bottom of the band and goes up along the eigen energy versus the *k* curve giving

increasing velocities and hence, conductivities until it reaches the Brillouin zone edge as shown in Fig. 5.15.



Fig. 5.15 Schematic showing the variation of energy for variation of t

Here, the electron at the zone edge is "reflected" and starts to loose energy in its motion in the field. The *k*-direction of the electron changes sign as the electron passes through the zone edge representing oscillations in *k*-space and this oscillation in energy continues. Thus, instead of infinitely increasing conductivity due to increasing ballistic velocity, an oscillation or AC current is observed. This oscillation of energy has a maximum value corresponding to *k*-value of $2\pi/L$ for lattice potential of spatial period *L*. The time taken by the electron to cover a Brillouin zone can be estimated from the equation given by

$$\Delta k = \frac{eE\tau_B}{\hbar} = \frac{2\pi}{L}$$

$$\tau_B = \frac{2\pi\hbar}{eEL}$$
(5.32)

i.e.,

and therefore, the oscillation frequency is

$$\omega_B = 2\pi \cdot \frac{1}{\tau_B} = \frac{eEL}{\hbar}$$
(5.33)

This energy oscillation is equal to a periodic motion in real space. These energy oscillations of a particle like an electron confined in a periodic potential when a constant force is acting on it are called *Bloch* oscillations.

In order to observe Bloch oscillations, the time period τ_B has to be less than the scattering time τ_{SC} . If the periodic distance *L* of the lattice is increased, then it will be possible to observe Bloch oscillations. One way to implement this is to change lattice constant artificially by introducing another material of thin lower bandgap alternatively and realizing an artificial lattice called the superlattice with a larger periodicity, so that the distance in the *E-k* space is reduced and Bloch oscillations can be completed before the electron undergoes scattering.

When the external field is strong, the electrons tunnel between Bloch bands, and is called Zener tunneling. Successive Zener tunneling to higher bands will lead to decay because the band gaps decrease with increasing energy. Zener tunneling should be limited to two closely spaced bands well separated from the rest of the bands in order to observe the coupling between Bloch and Zener tunneling. Modulating every second period of the potential, the original Bloch band splits into two sub-bands. The dynamics of quantum particles then shows a coherent superposition of Bloch oscillations and Zener tunneling between the sub-bands, called Zener-Bloch oscillations.

The oscillation frequency is quite high and can easily be in the several terahertz regions. Amplifiers and oscillators operating at high-frequencies are built using the superlattices. The Bloch oscillator is a device, where a DC-biased superlattice is connected to a resonator whose resonant frequency is the desired alternating field frequency.

5.9 RESONANT TUNNELING

Resonant tunneling in semiconductor structures is an area of interest for both fundamental research and applications in high speed electronics. Tunneling is a quantum process in which a particle penetrates into and traverses a barrier region where barrier potential exceeds its total energy. According to classical physics, conservation of the energy would not permit this tunneling and the particle is totally reflected by the barrier. In quantum physics, the particle exhibits wave nature with the de Broglie wavelength, and hence, its probability density is spread out over both the classically allowed and classically forbidden parts of the path, provided potential barrier is finite and thin. Wave functions penetrate from allowed regions into adjoining forbidden region and dies off exponentially with distance, and this allows non-classical particle motions that can cross this sufficiently narrow barrier. It is shown in Fig. 5.16, where the particle, e.g. electron is described by plane waves which is partly reflected and partly transmitted by the potential barrier.



Fig. 5.16 Illustration of the tunneling

Resonant tunneling is distinguished from the simple tunneling process by the presence within the classically forbidden region of quasi-bound or metastable states of the tunneling particle. The most typical situation is the formation of the double barrier resonant tunneling, as shown in Fig. 5.17. In this case, a quantum well is sandwiched between two thin finite potential barriers. The finite and thin potential barrier results in the penetration of the wave function of the quantum well into the classically forbidden barrier region resulting in a nonzero probability of penetration of the particle into the barrier regions. Electrons or free charge carriers are able to tunnel into the next region; if in addition, there are unfilled allowed states in the region in which it is penetrating.

For electrons with an energy corresponding approximately to the virtual resonant energy level of the quantum well, the transmission coefficient is close to unity. In general, the extension of periodic rectangular potential (Kronig-Penney) to the double barrier resonant tunneling, offers a simple model of the behaviour of electrons. The Schrödinger equation is solved separately for each region of the structure and the wave functions and their derivatives are matched at the boundaries to arrive at the tunneling probabilities. Figure 5.17 (a) shows a typical potential profile for a resonant tunneling structure. Here, the double barrier structure has a quasibound ground state at energy E_0 . The level E_0 is close to the level in the quantum well formed within the double barrier region but it is broadened due to the escape lifetime. The broadening comes from the Heisenberg energy-time uncertainty. If the electrons coming from the left have energies close to E_0 , they are able to transmit through the structure.



Fig. 5.17 Schematic of operation of resonant tunneling devices (a) showing the energy band diagram for different bias voltages and (b) current-voltage characteristics

The operation of a resonant tunneling structure is understood conceptually by examining Fig. 5.17. In zero bias, at point A, no current flows through the structure since the allowed empty level in the well is not aligned with the energy of electrons coming from the left. At point B, applying the bias voltage lifts up the Fermi level on the left of the barrier and when the Fermi energy lines up with the quasi-bound state, a maximum amount of current flows through the structure. Further increasing the bias results in the structure of point C,

where the current through the structure decreases with increasing bias (negative resistance). When a larger bias is applied, it results in a strong thermionic emission current and thus, the current increases substantially as shown at point D.

One important feature of resonant tunneling through double barrier is the sharp peaks in the transmission probability (approaching unity) at the quasi bound states of the quantum well formed between the two barriers with associated large peak to valley ratios in the negative resistance region.

5.10 QUANTUM INTERFERENCE EFFECTS

In nano devices, especially at low temperatures, where device dimensions are less than the typical mean free path (~ 100 Å) of electrons, the electrons are able to move without scattering, and hence, remain phase correlated. As they are interacting with structures of the order of their de Broglie wavelength, they exhibit wave nature rather than particle nature. Therefore, this quantum confined electrons exhibit all wave nature phenomenon like interference normally associated with photons or electromagnetic waves. Interference exhibited by quantum confined electrons at low temperatures due to its phase coherent transport and dominant wave nature in nanodevices is called *quantum interference*.

Quantum interference research is applied in a growing number of applications, such as the superconducting quantum interference device (SQUID), quantum cryptography, and quantum computing. In typical experiments demonstrating the interference, a minimum of two phase correlated beams are made to superimpose, and depending on their phase difference, they combine constructively or destructively to create interference patterns. In the case of quantum interference devices as shown in Fig. 5.18, this involves creating two pathways of same length say x, for the quantum confined electrons to flow, with a possibility of independently controlling the propagation through these pathways, to realise low voltage digitally controlled switches and other similar devices.



Fig. 5.18 *Quantum interference structure*

In analogy with optical interference, the intensity of the electron waves at output location z, is given by

$$I_{\text{output}}(z) = \left|\psi_1(z) + \psi_2(z)\right|^2 = \left|Ae^{ik_1z} + Ae^{ik_2z}\right|^2 = 2A^2\cos(k_1 - k_2)z$$
(5.34)

where $\psi_1(z)$, $\psi_2(z)$ are the wave functions of matter waves of the electrons with propagation constants of k_1 and k_2 , respectively travelling through paths 1 and 2 and reaching the point *z*.

A schematic of a split-gate transistor exploiting quantum interference effects is shown in Fig. 5.19. Here, the

electrons propagate from the source to the drain under the two independently controlled gates in the 2-dimensional channel of AlGaAs/GaAs. The source and drain are doped sections that have free charge carriers and the middle undoped region creates a potential barrier preventing the flow of charge carriers. Resonant tunnelling of the electrons occurs in the region under the gates when appropriate external bias is applied and hence, reduction in the barrier height, thereby, creating a channel through which electrons can flow to the drain. The ungated region provides a potential barrier for electron transport, thereby, creating isolation between the two paths. If wavevectors of the electrons flowing through these two paths can be



Fig. 5.19 Schematic of split-gate transistor

altered, it can modulate the signal at the drain. This modulation can be done by using an electric bias to alter the kinetic energy of the electrons in one arm. Interference effects are thus caused by altering the gate bias. Application of appropriate bias E, changes E_F which can result in a change in the electron k-vector.

Superconducting quantum interference device

A quantum interference device realized using superconducting materials separated by a thin layer of normal material, called the superconducting quantum interference device (SQUID) is the most sensitive magnetic flux detector. Here, the bias is altered by the external magnetic field. The device, operating at cryogenic temperatures, with quantum-limited sensitivity has demonstrated magnetic field resolution of 10^{-17} Tesla. SQUID can be used to passively measure extremely small variations in magnetic flux like those associated with human brains and has far reaching applications in medical field.

5.11 MESOSCOPIC STRUCTURES: CONDUCTANCE FLUCTUATION AND COHERENT TRANSPORT

Mesoscopic structures are those which have dimensions in the order of few tens of atoms. As discussed in previous sections, this result in quantum confinement of a countably small density of electrons which exhibit wave nature and maintain phase coherence in the absence of scattering, which is generally ensured at low temperatures. Unlike macroscopic devices with high density of electrons and lack of phase information of the electrons due to the dominant scattering limited transport, in mesoscopic structures, any factor that can affect the phase coherence of the electron transport like a small change in electron numbers, presence of a scatterer, or presence of small thermal or magnetic fields can lead to drastic observable changes in its conductivity. The charge quantization and wave nature of electrons result in quantization of conductance of these structures and also associated effects like coulomb blockade. The *V-I* characteristics of these structures do not follow Ohm's law. The electron transport in the few electron low scattering structure is predominantly ballistic transport with quantized conductance.

5.12 CONDUCTIVITY OF METALLIC NANOWIRES

A structure which is extended in one direction and constrained in the other two directions in the order of nanometer is called quantum wire or nanowire. Nanowires can be used in computer memories and other computing devices due to the special properties of high performance electronic components. The variation of

5.18 Physics for Electronics Engineering and Information Science

electrical properties and optical anisotropy of one-dimensional metals such as thin wires are due to its size. Metallic nanowires like silver nanowires are used in applications such as flexible touch screen, solar cells and transparent electrodes.

The electrical resistivity of the medium is caused by the scattering of electron by impurities, during which electrons oscillate around their equilibrium position in a solid. The mean free path of a moving electron is the average length that the electron can travel freely before collision. The mean free path can be increased by reducing the number of impurities in a crystal or by lowering its temperature.

If radius of metallic nanowire is less than the mean free path, the conductivity is different from the bulk material. The conductivity of nanowire decreases as the radius of nanowire is reduced. This is due to scattering from the wire's surface and from grain boundaries. In the bulk metals, the Lorentz number is temperature dependent. Its value equals the Sommerfeld value $(2.44 \times 10^{-8} \text{ W} \cdot \Omega/\text{K}^2)$ at high temperatures and extreme low temperatures. For intermediate temperatures, the Lorentz number decreases with decreasing temperature. In nanomaterials, due to grain boundary-electron reflection and electron-phonon scattering, the Lorentz number does not obey the property of bulk metal.

When the grain size of nanocrystalline metallic wire is less than the electron mean free path, the grain boundary-electron and surface-electron scatterings are intensive. The energy of scattered electrons can be partly transferred across the grain boundary via electron-phonon scattering because phonons can be transported through the grain boundary more readily than electrons. This causes reduction in electrical and thermal conductivity. In case of nano crystalline metallic wires, the Lorentz number is larger than that of Sommerfeld value. The conductivity of electrons in the nanowires can be explained by particle-in-a-box model.

For metallic nanowires, a size-dependent value of conductivity has to be used. The resistivity ($\rho = 1/\sigma$) of rectangular cross-section nanowires can be represented by

$$\rho = \rho_0 \left\{ \frac{1}{3\left[\frac{1}{3} - \frac{\alpha}{2} + \alpha^2 - \alpha^3 \ln\left(1 + \frac{1}{\alpha}\right)\right]} + \frac{3}{8}C(1-p)\frac{1+AR}{AR}\frac{L_m}{\omega} \right\}$$
(5.35)

where $\alpha = \frac{L_m}{d} \frac{R_c}{1 - R_c}$.

Here, σ is the conductivity of nanowire, ρ_0 is the resistivity of bulk solid, w is the wire width, AR is the aspect ratio (wire height divided by wire width, d) p is the specularity parameter (relating to reflection from the wire surface), R_c is the reflectivity coefficient, and C is a constant ≈ 1.2 and L_m is the mean free path. The first term is related to grain boundary electron scattering, and the second term corresponds to wire surface - electron scattering. The experimental results lead to p = 0.3 - 0.5 and $R_c = 0.2 - 0.3$. Here, both p and R_c values take the range between 0 and 1.

The V-I characteristic of a 30 nm radius, $2.4 \,\mu$ m long single-crystalline copper nanowire is shown in Fig. 5.20 at room temperature. This resistance is approximately 10 times the value of bulk copper.





5.13 BALLISTIC TRANSPORT

When the length of the wire becomes extremely small compared to the mean free path, an interesting phenomena called ballistic transport is observed, which is not found in macroscopic conductors. When the structure sizes are smaller than the mean free path of the electron, the motion of the electron in an applied electric field *E* is similar to that in an undamped system as there is no scattering. If the length of the wire *L* is reduced much less than mean free path L_m i.e., $L << L_m$, no collisions take place and hence, the collision-based model is not valid.

The two types of collisions of electrons that occur in a material are as follows:

- First one is elastic collision, in which, when an electron collides with an object there is no change in energy.
- (ii) Second type is elastic collision where the energy of the electron changes i.e., the total energy is conserved.

The following definitions define the different length scale of a material.

- 1. *L* is the length of the conductor.
- 2. L_m is the mean free path and is equal to the length that the electron can travel before having an elastic collision.
- 3. L_{ϕ} is the length over which an electron can travel before having an inelastic collision.

This is also called the phase-coherence length, since it is the length over which an electron wavefunction retains its coherence (i.e., retains its phase memory).

Over the phase-coherence length, the phase of the wavefunction evolves smoothly. The wavefunction, assuming a time-independent potential *E* is given by $\Psi(r,t) = \Psi(r)e^{-iEt/\hbar}$. Elastic collisions do not change phase coherence, but inelastic collisions make changes in phase coherence. These inelastic collisions are called dephasing events, which will destroy interference effects, including electrons interfering with themselves.

The electron transport is divided into two regions as follows:

- 1. When $L \gg L_{\phi}$, classical transport L_m is observed and applied to macroscopic dimension. The semiclassical or even classical models work well in this case.
- 2. When $L \ll L_m$, ballistic transport L_{ϕ} is observed. Ballistic transport occurs over very small length scales, where there is no momentum or phase relaxation. In a ballistic material, the electron's wave function can be obtained from Schrödinger's equation.

5.13.1 Ballistic Transport Model

A nanowire having small width, diameter or cross section in the order of Fermi wavelength (λ_F) and length of wire $L \ll L_m, L_{\phi}$, oriented along the *x* coordinate, extending from x = 0 to *L*, is considered. At both ends of the wire, there are reservoirs of electrons, which are simply macroscopic metal contacts that act as infinite sources and sinks for electrons, as shown in Fig. 5.21.

If the reservoirs have dimensions $D \times D \times D$, then the electrons in the reservoir have energy given by

$$E_n = \frac{\hbar^2 \pi^2}{2mD^2} n^2$$
(5.36)

where $n = 1, 2, 3, ..., \hbar$ is the reduced Planck's constant, *m* is the mass of free electron.

When *D* is sufficiently large, the energy states in the reservoir form essentially a continuum, as in a classical model. This continuum is the infinite source and sinks for electrons.



Fig. 5.21 Ballistic Channel connecting two electron reservoirs

5.13.2 Ballistic Channel and Subbands

The quantum mechanical picture is considered for ballistic transport between the reservoirs. The electrons in the wire will have energy of

$$E = \frac{\hbar^2}{2m_e^*} (k_x^2 + k_{y,n_y}^2 + k_{z,n_z}^2)$$

= $E_{n_y n_z} + \frac{\hbar^2}{2m_e^*} k_x^2$ (5.37)

where k_x is the longitudinal wavenumber, and is continuous because electrons are free along the x coordinate and the discrete indices n_y and n_z are subband indices. The subband energy levels are given by

$$E_{n_y n_z} = \frac{\hbar^2}{2m_e^*} (k_{y, n_y}^2 + k_{z, n_z}^2)$$
(5.38)

For an infinite confining potential (hard-wall case) and a rectangular cross-section $(L_y \times L_z)$ wire, the subband energy is

$$E_{n_y n_z} = \frac{\hbar^2}{2m_e^*} \left(\left(\frac{n_y \pi}{L_y} \right)^2 + \left(\frac{n_z \pi}{L_z} \right)^2 \right)$$

where $n_v, n_z = 1, 2, 3...$

When the quantum wire with a square cross section is assumed such that $L_y = L_z = L$, the hard-wall subband energy levels are given by

$$E_n = \frac{\hbar^2 \pi^2}{2m_e^* \omega^2} n^2$$
(5.39)

Then, the number of subbands (also called the number of electron channels, or modes) at Fermi energy or below the Fermi energy is given by

$$N = \sqrt{\frac{E_F 2m_e^* L^2}{\hbar^2 \pi^2}} = \frac{L}{\hbar \pi} \sqrt{E_f 2m_e^*} = \frac{Lk_F}{\pi} = \frac{2L}{\lambda_F}$$
(5.40)

where the usual E - k relationship for electrons, given by

$$E = \frac{\hbar^2 k^2}{2m_e^*}$$
(5.41)

Therefore, the width of the wire increases and hence, the number of electron channels also increases, so that the wire gains a new channel (mode) each time, the width becomes equal to an integral multiple of a half-Fermi wavelength. It is similar to electromagnetic wave propagation in a hollow conducting waveguide.

Applications of Ballistic Transport

The following are the applications of ballistic transport.

- 1. It is used in ultra-short-channel semiconductor FETs and carbon nanotube transistors.
- 2. Short interconnects.

5.14 QUANTUM RESISTANCE AND CONDUCTANCE

When electronic mean free path L_m of a wire is larger than the wire's length L, the wire behaves like an electron wave-guide and each wave-guide mode or conduction channel-contributes exactly an amount called quantum conductance to the total conductance of the wire.

Unlike the situation in bulk materials, in the nanostructures, the current given by $I = n \cdot e \cdot v_k$ does not increase linearly with applied voltage, as k values are quantized. Let δI be a small change in current in the presence of external drive potential of δV . Here, δI can be written as $\delta I = \delta n \cdot e \cdot v_k$ where δn is the change in number of electrons and v_k is the velocity of electron. As per quantum mechanics, velocity of electrons are given from the *E-k* relationship $v_k = \frac{1}{\hbar} \frac{\delta E}{\delta k}$. As change in electrons $\delta n = \frac{\delta n}{\delta E} (e \delta V)$, the expression for change in

current becomes $\delta I = \frac{e^2}{\hbar} \frac{\delta n}{\delta k} \delta V$; and therefore, the quantum conductance per channel, $G_0 = \frac{\delta I}{\delta V} = \frac{e^2}{\hbar} \frac{\delta n}{\delta k}$.

Here, the conductance is a constant which is independent of length L.

Taking an example of confinement in one dimension, the number of electron states per k-space dn/dk is $1/\pi$, giving conductance value as $G_0 = (e^2/\pi\hbar) = (2e^2/\hbar)$, i.e., the conductance is quantized. As the width of the conductance channel increases, the number of subbands also increases and the quantum conductance is given by

$$G = \frac{2e^2}{h}N\tag{5.42}$$

This equation is called *Landauer formula*. Here, N is the number of quantum conductance channels.

Substituting Eqn. (5.40) in Eqn. (5.42), we have

$$G = \frac{2e^2}{h} \times \frac{2L}{\lambda_F}$$
(5.43)

where *L* is quantum wire cross section and λ_F is Fermi wavelength. For every unit rise in potential, the current increases by a fixed amount of $(2e^2/h)$. The *V-I* characteristics of these structures shown in Fig. 5.22 are therefore not linear but shows increase in quantized steps of $(2e^2/h)$. This phenomenon is called *quantum conductance*. Each jump in the conductance indicates one additional one dimensional pathway for transport within the nanostructure, and hence, the conductance fluctuation occurs.

Quantum resistance of each channel is

$$R_0 = \frac{1}{G_0} = \frac{h}{2e^2} = 12.9 \,\mathrm{k}\Omega$$

where $G_0 = 2e^2/h$ and the total resistance is

$$R = \frac{1}{G} = \frac{h}{N2e^2} = \frac{12.9}{N} k\Omega$$

As the number of channel increases, the conductance increases and hence, the resistance decreases.

The quantum conductance is shown in Fig. 5.22 for low temperature experiments. The material between the source and drain is a two dimensional electron gas and it is connected to a gate voltage V_G . When $V_G = 0$ the



Fig. 5.22 *Quantum conductance with respect to gate* $voltage V_G$

channel is open and the number of electronic channels is determined by the width of the physical channel. In this case, more electron subbands are obtained. When the gate voltage is increased, the width is narrowed and the subbands are decreased. This leads to decrease in conductance and increase in resistance. This forms a step like behavior as shown in Fig. 5.22.

The phase coherence is highly sensitive to external parameters or any change in internal structure. The fluctuations in the conductance in time, as a function of parameters like temperature or doping in nanodevices are being used to estimate the phase coherence lengths of each of these effects and served as an excellent measure for these changes. Hence, conductance fluctuations refer to the change in the conductance with time, or the temporal fluctuations of conductance, of a quantum device under constant bias, due to temporary loss in phase coherence induced by external parameters like temperature, or scattering from internal structural changes.

5.15 COULOMB BLOCKADE EFFECTS

In the nanoelectronic devices like quantum dot, electrons inside the device will create a strong coulomb repulsion and prevent other electrons from flowing through it. In large systems, the repulsion is negligible, and in very small systems, repulsion is significant due to electron charging energy. The increased resistance observed at low temperatures in the semiconductor devices is called Coulomb blockade. The Coulomb blockade is a tunneling effect, and its main feature is the total suppression of the current in a finite interval of external bias voltage. These small devices will not follow the Ohm's law.

A nano capacitor generally consists of a thin insulator between two conducting electrodes. The structure is like a tunnel junction consists of a thin insulating barrier between two conducting electrodes. If the device dimensions are such that the capacitance given by $C = \frac{\varepsilon A}{d} = \frac{\Delta q}{\Delta V}$ is very small (A is the area of the capacitor and d is the thickness of the insulating barrier), to create a significant change in capacitance even by addition of a single charge, the energy required is $E = \frac{1}{2}CV^2 = \frac{q^2}{2C}$. Hence, the voltage required is $V = \frac{q}{C}$. This

value is nearly 80 mV for a capacitance change of atto farad (10^{-18} F) . This implies that tunnelling current will flow only when a sufficient large voltage exists across the capacitor.

If this energy is supplied thermally, equivalent temperature required will be

$$T_{\rm th} = \frac{E}{k_B} = 928 \text{ K/C} \text{ (in atto Farad),}$$

where k_B is Boltzmann's constant.

Therefore, if temperature is kept low in very small capacitance tunnel junctions, the presence of a Coulomb blockade ensures no current flow until the voltage reaches a threshold value determined by the charging energy. If the device dimensions are such that capacitance is 10^{-18} F, then the current does not flow unless temperature is 928 K. On the other hand, if dimensions are adjusted to give a capacitance of say 10^{-16} F, then $T_{\rm th}$ reduces to 10 K and charging voltage requirement drops to 1mV. Thus, different applications can be realised through control of device dimensions. The *V-I* characteristics representing the Coulomb blockade is shown in Fig. 5.23. As the temperature rises, the flow of carrier increases and the Coulomb blockade effect vanishes.



Fig. 5.23 Coulomb blockade effect in nano-scale capacitor

5.16 SINGLE ELECTRON PHENOMENA AND SINGLE ELECTRON TRANSISTOR

The incessant requirement of technologies that can provide higher chip density, low power consumption and high speed, is driving the microelectronic industry towards pursuing new options for realisation of electronic switching through manipulation of a single electron or even spin of a single electron. In this context, quantum dot structures whose dimensions can be designed to realise quantum confinement of one or a few electron are gaining lot of importance. A quantum device that performs switching action by the removal or injection of a single electron is called a single electron transistor (SET). The single electron transistor works using a combination of coulomb blockade and resonant tunnelling effects.

Quantum dots with sizes in the order of few hundred nanometres contain only a discrete number of electrons. Addition of a single electron to this system requires an electrostatic charging energy due to the repulsive electron-electron interactions. Coulomb blockade is the phenomenon of inhibition of tunnelling of an electron into a quantum dot at low temperatures and for applied voltages below the charging energy of the structure. The resonant tunnelling employed in the SET refers to controlling the bias of the quantum dot (QD) to supply this charging energy to facilitate tunnelling of a single electron into the available quantum state of the QD.

There are many variations in the structure of the single-electron transistor. The main components of the single electron transistor are shown in Fig. 5.24. The quantum dot is connected to the source and drain through thin insulator regions which possess larger bandgaps when compared to the QD. The source and drain are generally doped semiconductors and the insulator regions could be undoped regions of thickness of few tens of nanometres. The gate terminal provides electrostatic or capacitive coupling. The source-drain conductance, G is zero for most gate voltages, because of Coulomb blockade. Adding one electron requires

the supply of Coulomb energy i.e., $q^2/2C$ where q is the charge of electron and C is the capacitance across the QD and the difference in energy ΔE between filled and empty quantum levels of the quantum dot. If a second electron is added to the same quantum level, ΔE is not required and only the Coulomb energy is needed. The electrostatically influenced electrons traveling between the source and the drain terminals need to tunnel through two junctions (barriers). The QD island is charged and discharged as the electrons cross it, and the relative energies of the island containing zero or one extra electron depends on the gate voltage. Thus, the charge of the island changes by a quantized amount q.



Fig. 5.24 *Components of the single electron transistor (SET)*



Fig. 5.25 Two-step tunnelling process in SET (a) Unbiased (b) Biased (c) Biased with V_G (Left barrier lowered) (d) Biased and (e) Biased with V_G (Right barrier lowered)

Under no bias condition, there will be no movement of electron from source to drain as their Fermi levels are at the same level and there is a difference between the quantum level of QD and that of the source and drain, apart from coulomb blockade in QD. This situation is shown in Fig. 5.25(a). Now, if a bias voltage is applied across drain and source (V_{sd}), then the Fermi level of source, S will go up and the Fermi level of drain,
D will go down. No tunnelling occurs due to coulomb blockade. This situation is shown in Fig. 5.25(b). When gate voltage is applied such that the coulomb energy $q^2/2C$ is supplied, barrier gets lowered and tunnelling occurs and an electron gets added as shown in Fig. 5.25(c). In this case, electron will get transferred from source to QD. When a gate voltage, V_G is reversed, resonant tunnelling from QD to the drain region will occur as shown in Fig. 5.25(e). Thus, the gate controls tunnelling of one electron from source to drain. This is the working principle of SET. Each time the electron is transferred from source to the drain, there will be a current peak indicating charge transfer. Therefore, the current output oscillates with each transfer of charge and is said to exhibit coulomb oscillations.

5.16.1 Types of SET

There are different types of single electron transistor (SET) including carbon nanotube SET, semiconductor nanowire SET, molecular SET

(a) Carbon Nanotube SET

CNT single-electron transistor (CNTSET) is an attractive device for obtaining high sensitivity and low power consumption since CNT has nano dimension. It has many potential applications in next-generation devices, such as highly sensitive sensors, memories, and quantum information devices. Typical SET has a unique current oscillation with significantly narrow peaks, i.e., Coulomb oscillation. In the CNTSET with single-electron charge storages, the current peak in Coulomb oscillation (Coulomb peak) will shift in position after injecting a single electron into the floating gate. Digital memory operations will be obtained in the CNTSET with single electron charge storages, since charged carriers can be detected from the shift in the narrow peak of the SET. Therefore, ultra low power consumption and completely digital operation can be achieved by fabricating CNTSET with single-electron charge storages. The typical structure of CNTSET is shown in Fig. 5.26.



Fig. 5.26 Schematic image of CNTSET with single electron charge storages

(b) Semiconductor Nanowire SET

The nanowire based SET is fabricated by Silicon Metal Oxide Semicondutor (Si-MOS) technology. Each transistor consists of a Si nanowire channel and three fine gates that cross over it. When a negative voltage is applied to the gate, the electrostatic barrier is formed in the Si nanowire channel. This is exactly the switch-OFF mechanism of a MOSFET. The category of such devices with tunable electrostatic potential are called Si nanowire charge-coupled device (CCD), which is also an emerging research area. The conventional metal-based SET pumps with fixed tunnel barriers exhibit the current of few pA due to the maximum frequency of a few tens of MHz. The nanowire SET devices operate upto 100MHz and show small but periodic CB oscillations of SET transistors at low temperature. Figure 5.27 (a) and (b) show the schematic top-view and the cross-sectional view of the semiconductor nanowire SET.



Fig. 5.27 SET transistor using Si nanowire MOSFET (a) schematic top view and (b) cross-sectional view

(c) Molecular SET

Electron transport properties of individual molecules have received attention due to the introduction of SET devices, which allow probing electronic, vibrational or magnetic excitations in an individual molecule. In a three-terminal molecular SET, the molecule is situated between the source and drain leads with an insulated gate electrode underneath as shown in Fig. 5.28. The insulating ligand on the periphery of the molecule acts as an isolating barrier and hence, the coupling between the molecule and the electrode is capacitive, and the magnitude is dependent on the molecule or lead distance and the ligand composition. The current can flow between the source and drain leads via a sequential tunnelling process through the molecular charge levels, which the gate electrode is used to tune. The conduction through a molecular SET occurs only when a molecular electronic level lies between the Fermi energies of the leads.



Fig. 5.28 Schematic of molecular SET

5.16.2 Advantages of SET

The advantages of SET are as follows:

- (i) Low energy consumption
- (ii) High sensitivity
- (iii) Compact size
- (iv) High operating speed
- (v) Simplified circuit
- (vi) Feature of reproducibility
- (vii) Simple principle of operation
- (viii) Straight forward co-integration with traditional complementary metal oxide semiconductor (CMOS) circuits.

5.16.3 Disadvantages of SET

The disadvantages of SET are as follows:

- (i) Integration of SET in a large scale: To operate SET at room temperature, large quantities of mono dispersed nano particles less than 10 nm in diameter are required. But, it is very hard to fabricate large quantities of SET by traditional optical lithography and semiconducting process.
- (ii) It is difficult to link SET with the outside circuit.
- (iii) Fabrication technology are complicated for SET

5.17 MAGNETIC SEMICONDUCTORS

Magnetic semiconductors are semiconductor materials that exhibit both semiconductor and magnetic properties. EuO and $CdCr_2S_4$ are some examples of magnetic semiconductors that have been available from the early 1960s. The low Curie temperature and poor reproducibility of these magnetic semiconductors have limited their applications. The diluted magnetic semiconductors (DMS) obtained by doping magnetic atoms into nonmagnetic semiconductor host material give good applications with coexistence of magnetics semiconductors. The other category includes oxides (ZnO, TeO, etc) doped with ferro magnetic material. Of these, ZnO based DMS is of special importance as it combines transparency in the visible with conductivity and piezoelectricity and the addition of magnetic spin can lead to development of novel devices. The schematic of magnetic semiconductors is shown in Fig. 5.29.



Fig. 5.29 Schematic of magnetic semiconductors (a) non-magnetic, (b) magnetic and (c) diluted

5.18 SPINTRONICS

Typical electronic devices work on the principle of transport and manipulation of charge carriers. The new technology, spintronics, is based on manipulation of the spin and charge of the carriers in semiconductors or semiconductor hetero structures. These spin-based devices are expected to create and manipulate spin-polarized electrons in a host semiconductor. In order to achieve an operational device, the electrons must be spin polarized and their polarization largely preserved as they travel through the semiconductor material. One advantage of transport of spin polarized carriers is that they can maintain their spin state even for transport distances of hundreds of microns and normal scattering phenomenon does not change their spin state.

Higher integration levels in smaller device is possible because of the special properties of spintronic devices like the high energy efficiency and generation of less heat in operation than any other semiconductor devices. The three main requirements for spintronics are:

(i) efficient spin injection,

- (ii) slow spin relaxation and
- (iii) accurate and reliable spin detection

Under equilibrium, the density of spin-up and spin-down states are the same. To realize spin-polarized states, there has to be a mechanism to produce non-equilibrium distribution. One method of ensuring injection of

carriers with specific spin orientation into the semiconductor is to inject the carriers through ferromagnetic or DMS contacts into the semiconductors as shown in Fig. 5.30. Another method is to employ optical manipulation of the spin state of the charge carriers in DMS, through exposure to circularly polarized laser beams. This method can tailor and manipulate its spin transport properties. Once these spin oriented charges are injected into the device, they maintain their spin states during their transport through the short device lengths provided they are not subjected to external magnetic fields.



Fig. 5.30 Schematic of Spin transistor with spin injection through Fe/DMS contacts

Spin relaxation or spin dephasing can occur through spin-orbit coupling or through coupling to nuclear spins but is low in confined structures like quantum dots giving lifetimes of 100 ns to 1 ms. Methods of manipulating the spin states include exposure to magnetic fields, exposure to electric fields which influences the spin-orbit coupling and with exposure to time dependent optical fields. Spin-up and spin-down states of the electrons have different propagation constants i.e., k values even when their energies are the same and this leads to different phase delays for these two groups of electrons. Therefore, quantum interference devices can be realized using these spin-up and spin-down states or through control of k vector through appropriate gate bias for the same spin-state electrons.

Detection schemes for spin-states can be through (i) the resistivity change in the ferromagnet-non-magnetic material interface, (ii) spin induced faraday rotation of incident light beam, (iii) transmission through the nanostructure when subjected to external magnetic field or (iv) anomalous Hall effect (Hall effect observed in the absence of external magnetic field) induced by the current in spin-polarized material.

5.18.1 Spintronic Devices

Spin-orbit coupling plays a major role in determining the lifetime of the spin-polarization in the spintronic device. The spin relaxation rate in vacuum is of order 10^{-6} s whereas in a semiconductor, it is 10^{-3} s. The advantage of spin-based devices is that they are non-volatile compared with charge-based devices. Some of the important spintronic devices are Giant Magnetoresistance Devices (GMR), Magnetic Tunnel Junction (MTJ), magnetic spin transistor.

5.18.2 Giant Magnetoresistance Devices

Giant magnetoresistance (GMR) is a quantum mechanical magneto resistance effect observed in thin film structures composed of alternating ferromagnetic (FM) and nonmagnetic (NM) layers. The giant magnetoresistance effect is the change of electric conductivity in a system of metallic layers when an external magnetic field changes the magnetization of the ferromagnetic layers relative to each other. A parallel alignment has usually a lower resistance than an anti-parallel alignment. The GMR effect shows significant reduction of the electrical resistance of the multilayer in the presence of a magnetic field. In Fe/Cr and Co/Cu multilayers, the magnitude of GMR can be higher than 100% at low temperatures. The structure of GMR is as shown in Fig. 5.31.



Types of GMR

The GMR structure has different types that includes multilayer GMR, spin valve GMR, pseudo-spin GMR, granular GMR

Multilayer GMR: In multilayer GMR, two or more ferromagnetic layers are separated by a very thin (about 1 nm) non-ferromagnetic spacer (e.g. Fe/Cr/Fe). The electrical resistance of the device is normally higher in the anti-parallel case and the difference can reach more than 10% at room temperature.

Spin valve GMR: In spin valve GMR, two ferromagnetic layers are separated by a thin non-ferromagnetic spacer (~3 nm), but without coupling mechanism of nuclear magnetic moments. If the coercive fields of the two ferromagnetic electrodes are different, it is possible to switch them independently. Therefore, parallel and anti-parallel alignment can be achieved, and normally, the resistance is higher in the anti-parallel case. This device is also sometimes called a spin valve.

Pseudo-spin GMR: Pseudo-spin valve devices are very similar to the spin valve structures. The significant difference is the coercivities of the ferromagnetic layers. In a pseudo-spin valve structure, a soft magnet will be used for one layer whereas a hard ferromagnet will be used for the other. This allows the applied field to flip the magnetization of one layer before the other, thus providing the same anti-ferromagnetic affect that is required for GMR devices.

Granular GMR: Granular GMR is an effect that occurs in solid precipitates of a magnetic material in a non-magnetic matrix.

GMR Electrical Characteristics

A typical GMR sensor output is shown in Fig. 5.32. As the diagram shows an omnipolar response, it has the same change in resistance for a directionally positive or directionally negative magnetic field.



Fig. 5.32 Typical GMR sensor output

Applications of GMR

- (i) The largest technological application of GMR is in the data storage industry. IBM was first to put on the market with hard disks based on GMR technology, and nowadays, all disk drives make use of this technology.
- (ii) The GMR read head sensor in a hard disk is built using a spin valve. Spin valve resistance demonstrates a steep change in the small field range close to H = 0. As the magnetic bits on the hard drive pass under the read head, the magnetic alignment of the sensing layer in the spin valve changes resulting in the resistance change.
- (iii) Other applications of GMR are as diverse as automotive sensors, solid-state compasses and non-volatile magnetic memories.

5.18.3 Magnetic Tunnel Junction (MTJ)

Two ferromagnets separated by a thin insulator form a Magnetic Tunnel Junction (MTJ) producing a magnetoresistive effect is called Tunnel Magnetic Resistance (TMR). If the insulating layer is thin in the order of a few nanometers, electrons can tunnel from one ferromagnet into the other. When the electrons are tunneling between two ferromagnetic metals, the magnitude of the tunneling current depends on the relative orientation of the magnetization of both electrodes as shown in Fig. 5.33. The current is dependent on the magnetization because the tunneling current is proportional to the product of the electrode density of states at the Fermi level, yielding separate majority and minority bands for electrons with opposite spins in ferromagnetic material. Assuming the spin conservation for the tunneling electrons, there are two parallel currents of spin-up and spin-down character.

The current between electrodes with the same magnetization direction should be higher than those with opposite magnetization. Here, the



Fig. 5.33 Schematic mechanism of tunneling magnetic resistance

electrons tunnel across the device depending on the availability of free states for the "right" spin direction. If the two magnetic layers are parallel, a majority of electrons in one will find enough states of similar orientation in the other, thereby, reducing the resistance of the device. However, if the electrodes are antiparallel, electrons with both spin directions will encounter a closed valve in either of the two plates, resulting in a higher total resistance. The TMR effect is larger than the GMR effect by about one order of magnitude, making it a good candidate for magnetic sensors and hard drives.

5.18.4 Magnetic Spin Transistor

An ordinary transistor is based on an N-P-N junction. The gate terminal controls the voltage across the P-type semiconductor depending on the direction of the applied voltage; free electrons are either attracted toward the gate or repelled away from it. The voltage controls the flow of current between the two N-type semiconductors, allowing the transistors to occupy both ON and OFF states. The problem with this transistor is its volatility. When power is shut OFF, the electrons in the P-type semiconductor diffuse and destroy their previous ON or OFF configuration. This is the reason why computers cannot be instantly turned ON and OFF. In a magnetic transistor, the magnetized ferromagnetic layers replace the role of N and P-type semiconductors.

In these devices, a non magnetic layer which is used for transmitting and controlling the spin polarized electrons from source to drain plays a crucial role. For functioning of this device, first the spins have to be injected from source into this non-magnetic layer and then transmitted to the collector. These non-magnetic layers are also called semimetals because they have very large spin diffusion lengths. The injected spins which are transmitted through this layer start processing before they reach the collector due to the spin-orbit coupling effect. In order to solve this problem, an electric field is applied perpendicularly to the plane of the film by depositing a gate electrode on the top to reduce the spin-orbit coupling effect. By controlling the gate voltage and polarity, the current in the collector can be modulated mimicking the MOSFET of the conventional electronics.

Figure 5.34 shows the circuit of spin polarized field effect transistor. When the gate voltage V_G is zero, the injected spins are transmitted through the nonmagnetic layer and starts precessing before they reach the collector, thereby, reducing the net spin polarization. When $V_G >> 0$, the precession of the electrons is controlled with electric field, thereby, allowing the spins to reach at the collector with the same polarization. Hence, the net spin polarization does not change.



Fig. 5.34 Spin polarized field effect transistor

5.19 CARBON NANOTUBES

Carbon can be bonded in different ways to construct structures with different properties. The carbon nanostructures have applications in medicine, gene, and drug delivery areas. The strength and flexibility of carbon nanotubes make them more usable than other nanoscale structures. The advancements in carbon nanostructure make carbon nanotubes have a significant role in nanotechnology engineering. The physical properties of CNT are derived from graphene, the allotrope of carbon. In graphene, carbon atoms are densely organized in a regular sp^2 bonded atomic-scale honeycomb (hexagonal) pattern, and this pattern is a basic structure for other sp^2 carbon bonded materials (allotropes). Carbon nanotube is theoretically distinct as a cylinder fabricated of rolled up graphene sheet. It can be divided into a single wall or multiple walls. Nanotube with single wall is described as single-wall carbon nanotube (SWCNT) and the one with more than one wall is multiwall carbon nanotube (MWCNT).

Depending on the way, the rectangular sheet of graphene is rolled to form a carbon nanotube as shown in Fig. 5.35, the SWCNT's are further classified into armchair, zig-zag or chiral single-walled nanotubes. The way the graphene sheet is wrapped is represented by a pair of lattice translational indices (n, m). Here, the integers *n* and *m* denote the number of unit vectors \vec{a}_1 and \vec{a}_2 along two directions in the honeycomb crystal lattice of graphene such that the chiral angle θ and length of chiral vector $\vec{C}_h = n\vec{a}_1 + m\vec{a}_2$ represent the peripheral dimensions of the CNT. The illustration of vectors for the armchair and zig-zag SWCNT can be seen in Fig. 5.35.



Fig. 5.35 Schematic diagram of a Graphene sheet rolled to form a carbon nanotube

Figure 5.36 shows the different structures of SWCNT. When m = n, the nanotubes are named arm chair nanotubes, when m = 0, the nanotubes are named zigzag nanotubes and when $m \neq 0$, the nanotubes are called chiral.

A multi wall carbon nanotube (MWCNT) can be visualized as a structure with concentric cylinders of increasing diameters with correspondingly larger hemispherical end caps terminating them at each end. MWCNTs are more commonly used in products than SWCNTs. The structure of a MWCNT is shown in the Fig. 5.37.



Fig. 5.37 Structure of a multi wall carbon nanotube

Synthesis of Carbon Nanotube

Carbon nanotubes are synthesized by different methods like arc method in inert gas atmosphere, arc method in liquid nitrogen, plasma arc method, laser evaporation, and chemical vapour deposition. The laser evaporation and chemical vapour deposition methods produce good quality of CNT with an increased yield. The synthesis of CNT employing the electric arc discharge is discussed below.

Arc Discharge Method

The arc discharge method uses a vacuum chamber and an inert gas supply as shown in Fig. 5.38. The production of CNTs in high yield depends on the uniformity of the plasma arc. In this method, Arc vaporization is used to produce CNTs by placing two carbon rods from end to end, separated by an approximate distance of 1 mm, placed in an enclosure. The arc discharge method with liquid nitrogen can also be used to produce CNTs. At low pressure, the enclosure is usually filled with inert gas. A high temperature discharge between two electrodes is created by applying a direct current of 50 to 100 A. The voltage is then discharged from the rod, and vapourizing the carbon present in the electrode. The surface of the carbon electrodes is vapourized by the high temperature discharge. This carbon vapour is allowed to cool and forms a small rodshaped electrode. Even single-wall nanotubes can be



Fig. 5.38 Synthesis of CNT by arc discharge method

produced, if the proper metal ions are introduced. When optimized, this method can turn roughly 30% of the carbon into carbon nanotubes.

5.20 PROPERTIES OF CNT

There are many applications of carbon nanotubes based on their unique properties like electrical conductivity, strength, elasticity, aspect ratio, field emission, thermal conductivity, thermal expansion and magnetic property. Some of them are explained below.

5.20.1 Electrical Conductivity

A metallic CNT is a highly conductive material. Chirality, the degree of twist of graphene sheet, determines the conductivity of CNT interconnects. Depending on the chiral indices, CNTs exhibit both metallic and semiconducting properties. The electrical conductivity of MWCNTs is quite complex as their inter-wall interactions distribute the current non-uniformly over individual tubes. However, an uniform distribution of current is observed across different parts of metallic SWCNT.

5.20.2 Strength and Elasticity

The carbon atoms of a single graphene sheet of graphite form a planar honeycomb lattice in which each atom is connected via a strong chemical bond to three neighboring atoms. Due to these strong bonds, the elastic modulus of graphite is largest of any known material. Hence, CNTs exhibit the high-strength fibers. SWCNTs are stiffer than steel, and are resistant to damage from physical forces. The tip of a nanotube will bend when it is pressed. When the force is removed, the tip returns to its original state. This property makes CNTs very useful as probe tips for very high-resolution scanning probe microscopy.

5.20.3 Thermal Conductivity and Expansion

Ultra-small SWNTs has *superconductivity* below 20K. The strong in-plane graphitic C-C bonds make CNTs exceptionally strong and stiff against axial strains. The almost zero in-plane thermal expansion but large inter-plane expansion of SWNTs implies strong in-plane coupling and high flexibility against non-axial strains. CNT has very high thermal conductivity.

5.20.4 Magnetic Property

CNT exhibits the magneto resistance phenomenon like the bulk materials. The CNT displays the magneto resistive effect at low temperature, i.e., the resistance of the CNT is changed by the application of DC magnetic field. The observed magneto resistance effect is negative. This is due to the decrease in resistance with increase in magnetic field. The decrease in resistance results in an increase in conductance. When a DC magnetic field is applied to the nanotubes, the conduction electron acquires a new energy level known as Landau level. This energy level lies very close to the Fermi level and hence, the electrical conductivity increases.

5.20.5 Aspect Ratio

CNT has high aspect ratio, i.e. a low CNT load is required when compared to other conductive additives to achieve similar electrical conductivity. The high aspect ratio of CNTs possesses unique electrical conductivity in comparison to the conventional additive materials such as chopped carbon fiber, carbon black, or stainless steel fiber.

5.20.6 Field emission

Field emission results from the tunneling of electrons from a metal tip into vacuum, under the application of a strong electric field. The small diameter and high aspect ratio of CNT is an advantage for field emission. Even for moderate voltages, a strong electric field develops at the free end of supported CNT because of their sharpness. The CNT based field emitters must be superior to conventional electron sources and find its applications, especially in flat-panel displays.

5.21 APPLICATIONS OF CNT

Because of their unique properties like electrical, magnetic and mechanical, CNT find applications in coatings, energy storage, electronic devices, textile, electromagnetic etc. Some of them are given below:

5.21.1 CNT based Air and Water Filtration

CNT based air and water filters can block the smallest particles and kill most bacteria.

- (a) Air pollution filter: CNT is one of the best materials for air filters because it possesses high adsorption capacity and large specific area. The conductance of CNT changes when polluted gas comes in its contact. This helps in detecting and filtering the polluted air. CNT membranes can successfully filter carbon dioxide from power plant emissions.
- (b) Water filter: CNT membranes can aid in filtration and reduce distillation costs by 75%. Since these tubes are thin, the small particles like water molecules can pass through them while blocking the larger particles such as the chloride ions present in salt. CNT has effective sorption capacity over the pH range of 7 to 10.

5.21.2 CNT based Fibers and Fabrics

Fibers spun of pure CNT are the new development along with CNT composite fibers, which have many applications including body and vehicle armor, transmission line cables, woven fabrics and textiles. CNT is also used to make textiles stain resistant.

- (a) Textiles-CNTs can produce water-proof and tear-resistant fabrics.
- (b) Body armor-CNT fibers are being used as combat jackets that are used to monitor the condition of the wearer and to provide protection from bullets.
- (c) Polyethylene-CNT fibers can be used as polyethylene which can increase the elastic modulus of the polymers by 30 %.
- (d) Fire protection-Thin layers of buckypaper can potentially protect the objects from fire. The dense, compact layer of CNT or carbon fibers in the form of buckypaper can efficiently reflect the heat.
- (e) Waterproof-CNTs can be used to prepare super hydrophobic cotton fabric by dip-coating approach. This approach is dependent on the chemical reactions caused by UV-activated nitrene solution. The solution is used to transform the cotton fabric surface from hydrophilic to super hydrophobic with an apparent water contact angle of 154°. The surface of the cotton fabric is covalently attached by CNT, the super hydrophobicity possesses high stability and chemical durability.

5.21.3 CNT based Structural Composites

CNT has superior mechanical properties like stiffness, toughness and strength which lead to a large number of applications, including advanced composites.

- (a) Concrete-CNT in concrete increases its tensile strength and halt crack propagation.
- (b) Bridge-CNT may be able to replace steel in suspension bridges.

5.21.4 CNT based Electronic applications

CNT can be fabricated as electrical conductors, semiconductors and insulators. Its applications include the following:

- (a) Buckypaper-Thin nanotube sheets are 250 times stronger and 10 times lighter than steel. They can be used as heat sink for chipboards, backlight for LCD screens, or Faraday cage to protect electrical devices/aeroplanes.
- (b) Light bulb filament-CNT can be used as alternative to tungsten filament in incandescent lamp
- (c) Magnets-A strong magnetic field can be generated using multi-wall CNT coated with magnetite
- (d) Solar cells-Germanium CNT diode exploits the photovoltaic effect. In some solar cells, nanotubes are used to replace the ITO (indium tin-oxide) to allow the light to pass through active layers and generate photocurrent.
- (e) Electromagnetic antenna-CNT can act as an antenna for radio and other electromagnetic devices due to its durability, light weight and conductive properties. The skin effect in CNT is negligible at high frequencies due to additional kinetic inductance which results in low power dissipation.
- (f) Loudspeaker-Loudspeakers can be manufactured from sheets of parallel CNT. Such a loudspeaker can generate sound similar to the sound of lightning with thunder.
- (g) Sensors-CNT based sensors can detect temperature, air pressure, chemical gases such as carbon monoxide, ammonia, molecular pressure, strain, etc. The operation of a CNT based sensor is primarily dependent on the generation of current/voltage. The electric current is generated by the flow of free

charged carrier induced in any material. This charge is typically modulated by the adsorption of a target on the CNT surface.

- (h) Electrical circuits-The unique electrical properties of CNT are useful for building electronic devices, such as CNT field-effect transistors (CNTFETs) and CNT diodes. CNT can be used to form a *PN* junction diode by chemical doping and polymer coating. These types of diodes can be used to form a computer chip. CNT diodes can potentially dissipate heat out of the computer chips due to their unique thermal transmission properties.
- (i) Interconnects Carbon nanotube has emerged as one of the most potential interconnect material solutions in current nanoscale regime. The higher current density of 1000 MA/sq-cm of an isolated CNT can eliminate the electromigration reliability concerns that plagues the current nanoscale copper interconnects. Therefore, CNT interconnects have advantages over copper in terms of crosstalk, delay and power dissipation.

5.21.5 CNT based Energy Storage

CNT has the intrinsic characteristics and is used as electrode in battery and capacitor. CNT has high surface area, good electrical conductivity, and its linear geometry makes its surface highly accessible to the electrolyte. CNT is an outstanding material for super capacitor electrode. CNT also has applications in a variety of fuel cell components, which make them useful as electrode catalyst supports in proton-exchange membrane fuel cells. It is used in gas diffusion layers, and in current collectors, because of its high electrical conductivity.

5.21.6 CNT based Biomedical Applications

The exploration of CNT in biomedical application has significant potential. A large part of the human body consists of biocompatible material carbon. Cells have been shown to grow on CNT and hence, they have no toxic effect. The cells also do not adhere to the CNTs, potentially giving rise to applications such as coatings for prosthetics, and anti-fouling coatings for ships. The functional ability to the sidewalls of CNTs leads to biomedical applications such as vascular stents, neuron growth and regeneration. A single strand of DNA can be bonded to a nanotube and can then be inserted into a cell.

Note: Quantum dot laser topic is discussed in Section 4.22 of Chapter 4.

TWO MARK QUESTIONS AND ANSWERS

1. Define nano electronic device.

Electronic devices with at least one dimension in nanometer scale (≤ 100 nm) are called nano electronic devices.

2. What are the methods to get Nanostructures?

There are two methods to get nanostructures. They are:

- (i) The top-down approach, where the bulk material has been selectively etched to realise nanodevices using advanced lithography techniques
- (ii) The bottom-up approach where nanostructures are grown through layer-by-layer molecular or atomic assembly

3. Define quantum confinement.

If the volume of the material is reduced to an extent where the energy levels appear as discrete levels within each energy band, the number of electrons or charges confined in these volumes decrease and becomes quantized and countable. This confinement, which results in few countable numbers of electrons/charges, is called quantum confinement. If the dimension is reduced to be half the de Broglie wavelength (λ_{dB}) of the charge carrier, then single electron confinement is possible. In practise, quantum confinement is observed even for confined dimensions as large as a few multiples of λ_{dB} .

4. Mention any four important applications of nanomaterials.

The applications of nanomaterials are as follows:

- (i) Nano-MEMS (Micro-Electro Mechanical System) are used in ICs, optical switches, pressure sensors, mass sensors etc.
- (ii) Recently nano-robots were designed, which are used to remove the damaged cancer cells and also to modify the neuron network in human body.
- (iii) Nano-dimensional photonic crystals and quantum electronic devices plays a vital role in the recently developed computers.
- (iv) Bio-sensitive nano-particles are used in the production of DNA-chips, bio-sensors etc.

5. How does the band gap of metal change with particle size?

In a metal, the quasi-continuous density of states in the valence and the conduction bands split into discrete electronic levels. The spacing between these levels and the band gap increases with decreasing particle size. When the particle size is less than de Broglie wavelength of electrons, the charge carrier can be treated as particle in a box.

6. What is Zener tunneling?

When the external field is strong, the electrons tunnel between Bloch bands, and is called Zener tunneling.

7. Define Bloch oscillations.

The energy oscillations of a particle like an electron confined in a periodic potential when a constant force is acting on it are called Bloch oscillations. As energy supplied to the system increases, the charge carrier moves higher in the *E*-*k* band and reaches a maximum at the band-edge and undergoes reflection at the zone boundaries reversing its propagation loosing energy to return back to the lowest energy of the *E*-*k* band. The frequency of Bloch oscillations when the electron confined in a dimension *L* is subjected to a field of strength *E* is eEL/\hbar .

8. Explain the phenomenon for Bloch Zener oscillations.

The dynamics of quantum particles in a periodic potential with an additional constant force shows a coherent superposition of Bloch oscillations and Zener tunnelling between the sub-bands is called Bloch–Zener oscillations.

9. What is resonant tunneling?

The electron transmission through a quantum structure isolated by thin insulating barriers is called resonant tunneling. For electrons with an energy corresponding approximately to the virtual resonant

energy level of the quantum well, the transmission coefficient is close to unity and the electron with this resonant energy can cross the double barrier without being reflected.

10. What do you understand by coherent electron transport?

Undamped motion of an electron preserving its phase coherence in the absence of scattering is called coherent electron transport. This requires device dimensions to be less that the mean free path of the electrons and low temperatures to avoid scattering from lattice vibrations. Electrons then exhibit ballistic transport and quantised conductance where conductance is independent of the length of the device.

11. Define Coulomb blockade effect.

The increased resistance observed at low temperatures in the semiconductor devices is called Coulomb blockade effect.

12. Define single electron phenomenon.

The study of quantum dot structures whose dimensions can be designed to realise quantum confinement of one or a few electrons is called single electron phenomenon

13. What is single electron transistor?

A quantum device that performs switching action by the removal or injection of a single electron is called a single electron transistor (SET). The single electron transistor works using a combination of coulomb blockade and resonant tunnelling effects.

14. Define Giant magnetoresistance (GMR).

Giant magnetoresistance (GMR) is a quantum mechanical magneto resistance effect observed in thin film structures composed of alternating ferromagnetic and nonmagnetic layers.

15. Define spintronics.

The study of the intrinsic spin of the electron and its associated magnetic moment, in addition to its fundamental electronic charge is called spintronics. Spintronics deals with the electric, magnetic and optical properties of solids by the manipulation of spin population and their dynamics.

16. What are the applications of spintronics?

Electron spin in metals are used for computer magnetic data storage to store, read and write the data. They are used to realise low power high speed switching devices like magnetic tunnel junctions, magnetic spin transistors, magnetic spin FETs and also as very sensitive sensors for magnetic fields.

17. Calculate the quantum resistance of a single channel device.

Solution $R_0 = \frac{1}{G_0} = \frac{h}{2e^2N} = \frac{6.626 \times 10^{-34}}{2 \times (1.602 \times 10^{-19})^2 \times 1} = 12.9 \text{ k}\Omega$

18. Calculate the quantum conductance of quantum device with N = 10.

Solution
$$G_0 = \frac{2e^2N}{h} = \frac{2 \times (1.602 \times 10^{-19})^2 \times 10}{6.626 \times 10^{-34}} = 7.727 \times 10^{-4}$$
 \heartsuit

5.40 Physics for Electronics Engineering and Information Science

19. What is meant by carbon nanotube?

Carbon nanotube (CNT) are molecular scale tubes of graphitic carbon with outstanding properties. It is the stiffest and strongest fibre found till date, with remarkable electronic properties and applications.

20. List out the various forms of CNT.

The various forms of CNT are (i) Arm chair structure, (ii) Zig-zag structure and (iii) Chiral structure.

21. Give the techniques adopted to synthesis the carbon nano-tubes.

The following are the techniques adopted to synthesis the carbon nano-tubes.

- (i) Chemical vapour deposition
- (ii) Carbon-arc technique and
- (iii) Pulsed laser deposition method

22. How can CNT be produced by chemical vapour deposition method?

In chemical vapour deposition method, hydro-carbon gas such as methane is heated upto 1100°C and is decomposed. As the gas decomposes, it produces carbon atoms. These carbon atoms (chemical vapours) are made to deposit over a cooler substrate, which contains iron as a catalyst to form carbon nano-tubes.

23. What are the properties of CNT?

The properties of CNT are as follows:

- (i) The strength of the sp^2 carbon-carbon bonds gives high hardness for carbon nanotubes.
- (ii) The Young's modulus of the nanotubes is approximately 5 times greater than steel.
- (iii) The tensile strength is around 50 times higher than steel.
- (iv) The electronic properties of carbon nanotubes are also extraordinary. Especially, the caison nanotubes can be obtained in the form of metallic or semiconducting based on its structural content.
- (v) Nanotubes have higher conductivities than that of copper.

24. What are the applications of carbon nanotubes?

The applications of carbon nanotubes are as follows:

- (i) Carbon nanotubes (CNTs) are used in battery electrodes, fuel cells, reinforcing fibres etc.
- (ii) CNTs are used in the development of flat panel displays for computer monitors and televisions.
- (iii) Light weight CNTs are also used in military and communication systems, for protecting computers and electronic devices.
- (iv) Semiconducting CNTs are used as switching devices.
- (v) Semiconducting CNTs are also used as chemical sensors to detect various gases.
- (vi) The carbon nanotubes are very light in weight, but they are very strong and hence, they are used in aerospace.

25. Carbon nanotubes are used in high resolution scanning probe microscopy. Why?

Single wall carbon nanotubes are stiffer than steel, and are very resistant to damage from physical forces. Pressing on the tip of the nanotube will cause it to bend, but without damage to the tip. When the force is removed, the nanotube returns to its original state. This property makes CNT very useful as probe tips for very high resolution scanning probe microscopy.

REVIEW QUESTIONS

- 1. What are the classifications of nanomaterials?
- 2. Define the density of states.
- 3. What is the assumption that is necessary for the evaluation of the density states?
- 4. Define a Fermi energy level.
- 5. Derive the expression for the density of states for different quantum confinements.
- 6. Discuss about bandgap of nanomaterials.
- 7. What are metallic nanowires?
- 8. Describe the conductivity of metallic nanowires.
- 9. When does ballistic transport take place?
- 10. Explain the term quantum of conductance.
- 11. Define quantum of resistance.
- 12. What is quantum mechanical tunneling?
- 13. What is size quantization effect?
- 14. How does the band gap of semiconductors change with particle size?
- 15. Show a schematic representation of a planar quantum dot.
- 16. Show a schematic energy diagram of an AlGaAs/GaAs quantum well.
- 17. Derive the density of states in 2D nanomaterials or quantum wells with neat diagram.
- 18. Derive the density of states in 1D nanomaterials or quantum wires.
- 19. Explain the density of states in oD nanomaterials or quantum dots with neat diagram.
- 20. If all sides of a cubical box are doubled, what happens to the ground-state energy of a particle in that box? (Answer: *E* reduced to *E*/4)
- 21. An electron is confined to a cubical box. For what box width will a transition from the first excited state to the ground state result in emission of a 950 nm infrared photon? (Answer: 930 pm)
- 22. What will be the frequency of Bloch oscillations for an electron confined in a nanodevice of dimension 100 nm when driven by a field of 10 V/m. (Answer: 0.24GHz)
- 23. Describe resonant tunneling.
- 24. Explain Coulomb blockade effect in nano-scale capacitor with neat diagram.
- 25. Explain the construction and working of single electron transistor.
- 26. How is the template used to obtain nanowires or nanorods?
- 27. How can we store the nanoparticles?
- 28. Why do nanostructured particles find potential applications?
- 29. What is meant by surface induced effect?
- 30. How are nanostructured particles used for health applications?
- 31. Give two advantages of using spin to carry information.

5.42 Physics for Electronics Engineering and Information Science

- 32. What are dilute magnetic semiconductors?
- 33. What are MRAMS?
- 34. Explain ordinary magneto resistance effect.
- 35. What is GMR?
- 36. Distinguish between GMR and CMR.
- 37. What is meant by sensor?
- 38. Mention few applications of nanomagnetic materials.
- 39. Discuss the physical properties of carbon nanotubes.
- 40. What is meant by carbon nanotube?
- 41. Differentiate multi-walled and single walled carbon nanotubes.
- 42. Explain the various applications of CNT.