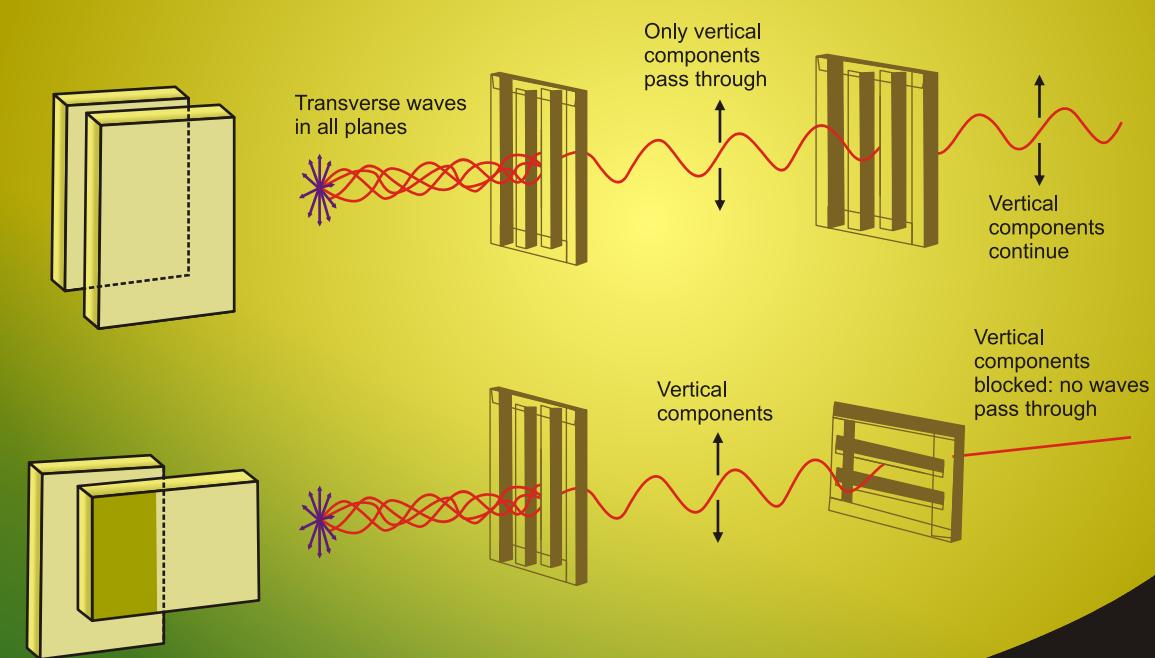


Eastern
Economy
Edition

ENGINEERING PHYSICS



G. ARULDHAS

ENGINEERING PHYSICS

G. Aruldhas

Formerly Professor and Head of Physics
& Dean Faculty of Science
University of Kerala

PHI Learning Private Limited

New Delhi-110001
2010

Rs. 250.00

ENGINEERING PHYSICS
G. Aruldas

© 2010 by PHI Learning Private Limited, New Delhi. All rights reserved. No part of this book may be reproduced in any form, by mimeograph or any other means, without permission in writing from the publisher.

ISBN-978-81-203-3916-3

The export rights of this book are vested solely with the publisher.

Published by Asoke K. Ghosh, PHI Learning Private Limited, M-97, Connaught Circus, New Delhi-110001 and Printed by Baba Barkha Nath Printers, Bahadurgarh, Haryana-124507.

*To
Myrtle and our children
Vinod & Anitha, Manoj & Bini, Ann & Suresh
for giving meaning to my life*



Contents

Preface

xvii

1. OSCILLATIONS AND WAVES	1-29
1.1 Harmonic Oscillation—The Basics	1
1.2 Simple Harmonic Motion (SHM)	1
1.2.1 Equation of Motion	2
1.2.2 Solution of Equation of Motion	2
1.2.3 Energy in SHM	4
1.3 Damped Harmonic Oscillation	5
1.4 Forced Oscillations and Resonance	8
1.5 Quality Factor	10
1.6 Types of Waves	11
1.6.1 Mechanical Waves	12
1.6.2 Electromagnetic Waves	12
1.6.3 Other Types of Waves	12
1.7 General Form of a Travelling Wave	12
1.8 Harmonic Wave	13
1.9 Wave Velocity and Particle Velocity	15
1.10 Energy Transport in Wave Motion	16
1.11 The General Wave Equation	17
1.11.1 One-dimensional Wave Equation	17
1.11.2 Three-dimensional Wave Equation	18
1.12 Solution of the One-dimensional Wave Equation	18
1.13 Solution of the Three-dimensional Wave Equation	19

vi *Contents*

1.14	Plane and Spherical Waves	21	
1.15	Transverse Vibrations of a Stretched String	21	
	<i>Solved Examples</i>	23	
	<i>Review Questions</i>	27	
	<i>Problems</i>	28	
2. ELECTROMAGNETIC THEORY			30–49
2.1	The Del Operator	30	
2.2	Gradient of a Scalar Function	31	
2.2.1	Physical Significance	31	
2.3	Divergence of a Vector	31	
2.3.1	Physical Significance	32	
2.4	Curl of a Vector	33	
2.4.1	Physical Significance	33	
2.4.2	Important Vector Identities	34	
2.4.3	Integral Theorems	34	
2.5	Conduction and Displacement Currents	35	
2.5.1	Conduction Current	35	
2.5.2	Displacement Current	36	
2.6	The Equation of Continuity for Time Varying Fields	36	
2.7	Maxwell's Equations	37	
2.7.1	Maxwell's Electric Flux Equation	37	
2.7.2	Maxwell's Magnetic Flux Equation	38	
2.7.3	Maxwell's Equation for Induced Electric Field	38	
2.7.4	Maxwell's Equation for Induced Magnetic Field	38	
2.8	Electromagnetic Waves in Dielectric Medium	40	
2.9	Transverse Nature of Electromagnetic Waves	41	
2.10	Relation between E and H	42	
2.11	Poynting's Theorem	44	
	<i>Solved Examples</i>	46	
	<i>Review Questions</i>	48	
	<i>Problems</i>	49	
3. INTERFERENCE			50–77
3.1	Superposition of Waves	50	
3.1.1	Analytical Treatment	51	
3.2	Young's Double Slit Experiment	53	
3.2.1	Experimental Arrangement	53	
3.2.2	Expression for Bandwidth β	53	
3.2.3	Conditions for Sustained Interference Pattern	55	
3.3	White Light Fringes	55	
3.4	Coherence	55	
3.5	Thickness of Transparent Sheet	56	
3.5.1	Expression for Thickness of Sheet	56	

3.6	Methods to Produce Interference Pattern	58
3.7	Interference in Thin Films	58
3.7.1	Reflected Light	58
3.7.2	Transmitted Light	60
3.8	Wedge Shaped Films	61
3.8.1	The Diameter of a Thin Wire	62
3.8.2	Testing the Planeness of Surfaces	62
3.9	Newton's Rings by Reflected Light	63
3.9.1	Experimental Arrangement	63
3.10	Newton's Rings with Liquid Film	65
3.10.1	Ring System in Transmitted Light	66
3.11	Michelson Interferometer	66
3.11.1	Principle	66
3.11.2	Experimental Arrangement	66
3.11.3	Working	67
3.11.4	Applications	67
3.12	Interference Filter	69
3.13	Antireflection Coatings	70
	<i>Solved Examples</i>	71
	<i>Review Questions</i>	74
	<i>Problems</i>	75
4.	DIFFRACTION	78–95
4.1	The Diffraction Phenomenon	78
4.2	Types of Diffraction	79
4.3	Fraunhofer Diffraction at a Single Slit	80
4.3.1	Experimental Arrangement	80
4.3.2	Positions of Maximum and Minimum Intensities	80
4.4	Diffraction at a Circular Aperture	82
4.5	Diffraction Grating	83
4.5.1	Transmission and Reflection Gratings	83
4.5.2	Experimental Arrangement	84
4.5.3	Grating Equation	85
4.5.4	Dispersive Power	86
4.6	Resolving Power	86
4.6.1	Rayleigh's Criterion	87
4.7	Resolving Power of a grating	87
4.8	Grating and Prism Spectra—Comparison	88
4.9	Resolving Power of a Telescope	89
4.10	Resolving Power of a Microscope	90
	<i>Solved Examples</i>	91
	<i>Review Questions</i>	94
	<i>Problems</i>	95

5. POLARIZATION	96–116
5.1 Introduction	96
5.2 Nature of Light	97
5.3 Types of Polarized Light	98
5.3.1 Linearly Polarized Light	98
5.3.2 Circularly Polarized Light	99
5.3.3 Elliptically Polarized Light	99
5.4 Types of Polarized Light—Analytical Treatment	99
5.5 Polarization by Reflection	101
5.5.1 Brewster's Law	101
5.6 Malus Law	102
5.7 Double Refraction	102
5.7.1 Uniaxial and Biaxial Crystals	103
5.8 Huygen's Theory of Double Refraction	104
5.9 Nicol Prism	104
5.9.1 Construction	104
5.9.2 As an Analyzer	105
5.10 Retardation Plates	106
5.11 Quarter Wave Plate	106
5.12 Half Wave Plate	106
5.13 Production of Different Types of Polarized Light	107
5.14 Polarization by Selective Absorption	107
5.14.1 Dichroism	107
5.14.2 Polaroid	107
5.15 Analysis of Light of Unknown Polarization	108
5.16 Kerr Effect	109
5.17 Faraday Effect	109
5.18 Photoelasticity	109
5.18.1 Basic Ideas	109
5.19 Induced Birefringence	110
5.20 Stress on a Photoelastic Material—An Analytical Treatment	110
5.20.1 Isoclinic and Isochromatic Fringes	112
5.21 Photoelastic Bench	112
<i>Solved Examples</i>	113
<i>Review Questions</i>	114
<i>Problems</i>	116
6. CRYSTALLOGRAPHY	117–141
6.1 Space Lattice, Basis, and Translation Vectors	117
6.2 Bravais Lattices	118
6.3 Point and Space Groups	120
6.4 Unit Cell	121
6.4.1 Primitive Unit Cell	120
6.4.2 Non-primitive Unit Cell	121

6.5	Lattice Points Per Unit Cell in Cubic Lattices	122	
6.5.1	Density and Number of Atoms Per Unit Cell	122	
6.6	Coordination Number	122	
6.7	Nearest Neighbour Distance	123	
6.8	Atomic Packing Factor	124	
6.9	Structure of Sodium Chloride	126	
6.10	Crystal Planes and Directions	127	
6.11	Miller Indices	127	
6.11.1	Significant Features of Miller Indices	128	
6.12	Interplanar Spacing	128	
6.13	X-ray Diffraction	130	
6.13.1	Bragg's Law	130	
6.13.2	X-ray Spectrometer	131	
6.13.3	Unit Cell Dimensions	132	
6.14	X-ray Diffraction—Laue's Method	132	
6.15	X-ray Diffraction—Powder Method	133	
	<i>Solved Examples</i>	134	
	<i>Review Questions</i>	140	
	<i>Problems</i>	141	
7.	SUPERCONDUCTIVITY		142–154
7.1	Introduction	142	
7.2	Magnetic Properties	143	
7.3	Meissner Effect	144	
7.4	Type-I and Type-II Superconductors	145	
7.5	The Isotope Effect	147	
7.6	BCS Theory	147	
7.6.1	Cooper Pairs	147	
7.6.2	Superconducting Energy Gap	148	
7.7	Magnetic Flux Quantization	148	
7.8	Josephson Junction	149	
7.9	High Temperature Superconductivity	150	
7.10	Applications	151	
7.10.1	Squid	151	
7.10.2	Generation and Transmission of Electricity	151	
7.10.3	Superconducting Magnets	151	
7.10.4	Other Applications	151	
	<i>Solved Examples</i>	152	
	<i>Review Questions</i>	153	
	<i>Problems</i>	154	
8.	SPECIAL THEORY OF RELATIVITY		155–179
8.1	Inertial and Noninertial Reference Frames	155	
8.2	Galilean Transformation	156	

x *Contents*

8.3	Electromagnetism and Galilean Transformation	157
8.3.1	Ether Hypothesis	157
8.4	Michelson-Morley Experiment	158
8.4.1	The Interferometer	158
8.4.2	The Experiment	159
8.4.3	Results	160
8.5	The Postulates of Special Theory of Relativity	161
8.6	Lorentz Transformation	161
8.7	Simultaneity	162
8.8	Time Dilation	163
8.9	Length Contraction	164
8.10	Velocity Addition	164
8.11	Mass in Relativity	166
8.12	Mass and Energy	169
8.13	Massless Particles	170
	<i>Solved Examples</i>	171
	<i>Review Questions</i>	177
	<i>Problems</i>	178

9. QUANTUM MECHANICS **180–209**

9.1	Breakdown of Classical Physics	180
9.1.1	Quantization of Matter and Charge	181
9.1.2	Photoelectric Effect	181
9.1.3	Specific Heat	181
9.1.4	Optical Spectra	181
9.1.5	Blackbody Radiation Curves	181
9.2	Planck's Quantum Theory	182
9.2.1	Planck's Hypothesis	182
9.2.2	Average Energy of an Oscillator	183
9.2.3	Planck's Radiation Formula	183
9.3	Dual Nature of Radiation	184
9.4	Early Models of the Atom	184
9.5	Matter Wave	185
9.5.1	De Broglie Hypothesis	185
9.5.2	Wave Packet	186
9.6	De Broglie Wavelength	186
9.7	The Davisson-Germer Experiment	187
9.8	Thomson's Experiment	188
9.9	Uncertainty Principle	189
9.10	Absence of Electron in the Nucleus	190
9.11	Uncertainty in Frequency of Light Emitted by an Atom	190
9.12	The Wave Function	190
9.13	Time Dependent Schrödinger Equation	191
9.13.1	Free Particle Wave Equation in One Dimension	191
9.13.2	Energy and Momentum Operators	193
9.13.3	General Wave Equation	193

9.14	Time Independent Schrödinger Equation	193
9.15	Probability Interpretation of the Wave Function	194
9.16	Normalization Condition	195
9.17	Operators	195
9.18	Hamiltonian Operator	196
9.19	Eigenvalues and Eigenfunctions of an Operator	196
9.19.1	Schrödinger Equation as an Eigenvalue Equation	197
9.20	Postulates of Quantum Mechanics	197
9.21	Particle in a One-dimensional Infinite Square Well	199
9.21.1	Energy Eigenvalues and Eigenfunctions	199
9.21.2	Probability Distribution	201
9.22	Quantum Mechanical Tunnelling	201
	<i>Solved Examples</i>	202
	<i>Review Questions</i>	207
	<i>Problems</i>	208

10. STATISTICAL MECHANICS 210–226

10.1	Statistical Methods	210
10.1.1	Macroscopic and Microscopic Systems	210
10.1.2	Macrostate and Microstate	211
10.1.3	Phase Space	211
10.1.4	Cells in Phase Space	212
10.2	Classical Statistics	212
10.3	Quantum Statistics	213
10.4	Basic Postulates of Quantum Statistics	213
10.5	Distribution Laws of the Different Statistics	214
10.5.1	Maxwell-Boltzmann Statistics	214
10.5.2	Bose-Einstein Statistics	214
10.5.3	Fermi-Dirac Statistics	215
10.6	Comparison of the Three Statistics	215
10.7	Density of States	216
10.7.1	Phase Volume and Eigenstates	216
10.7.2	Number of Eigenstates in an Energy Range	217
10.8	Planck's Blackbody Formula	218
10.9	Energy of a Fermi-Dirac System	219
	<i>Solved Examples</i>	222
	<i>Review Questions</i>	224
	<i>Problems</i>	225

11. LASERS 227–253

11.1	Absorption and Emission of Radiation	227
11.1.1	Absorption	227
11.1.2	Spontaneous Emission	228
11.1.3	Stimulated Emission	228
11.1.4	Radiative and Non-radiative Transitions	229

11.2	Principle of Laser	229
11.3	Einstein's Coefficients	229
11.4	Population Inversion	232
11.5	Basic Components of a Laser	232
11.5.1	Active Medium	233
11.5.2	Pumping Methods	233
11.5.3	Resonant Cavity	234
11.6	Metastable States	235
11.7	Three and Four Level Systems	235
11.7.1	Three Level System	235
11.7.2	Four Level System	236
11.8	Ruby Laser	236
11.8.1	Applications	238
11.9	He-Ne Laser	238
11.10	Carbon Dioxide Laser	241
11.11	Semiconductor Laser	242
11.12	Characteristics of Laser Light	244
11.12.1	Monochromaticity	244
11.12.2	Coherence	244
11.12.3	Directionality	244
11.12.4	Intensity	245
11.13	Applications of Lasers	245
11.13.1	Lasers in Industry	245
11.13.2	Lasers in Scientific Research	246
11.13.3	Lasers in the Medical Field	247
11.13.4	Other Applications	248
<i>Solved Examples</i>		249
<i>Review Questions</i>		252
<i>Problems</i>		253

12. HOLOGRAPHY AND FIBRE OPTICS **254–273**

12.1	Holography—An Introduction	254
12.2	Basic Description of Holography	255
12.2.1	Recording of Transmission Hologram	255
12.2.2	Reconstruction	255
12.3	Types of Holograms	256
12.3.1	Reflection Hologram	256
12.3.2	In-line and Off-axis Holography	256
12.3.3	Diffuse Illumination of the Object	256
12.3.4	Plane and Volume Holograms	257
12.4	Applications of Holography	257
12.4.1	Holographic Interferometry	257
12.4.2	Data Storage	257
12.4.3	Holographic Optical Element	257

12.4.4	Holographic Microscopy	258
12.4.5	Pattern Recognition	258
12.5	Optical Fibre	258
12.6	Light Propagation in Step Index Fibre	259
12.7	Numerical Aperture	260
12.8	Acceptance Angle and Acceptance Cone	261
12.9	Different Modes	261
12.10	Light Propagation in Variable Index Fibre	262
12.11	Single and Multimode Fibres	263
12.11.1	Salient Features of a Single Mode Fibre	263
12.11.2	Salient Features of Multimode Fibre	263
12.12	V-number of a Fibre	264
12.13	Fibre Losses and Transmission Windows	264
12.14	Optical Sources for Fibre Optics	265
12.15	Optical Communication	266
12.15.1	Communication System	266
12.15.2	Advantages of Optical Communication	267
12.16	Light Pipe	268
12.17	Fibre Optic Sensors	268
12.17.1	Intensity Modulated Sensor	269
12.17.2	Phase Sosnor	269
12.17.3	Sensors in Health Care	270
<i>Solved Examples</i>		271
<i>Review Questions</i>		272
<i>Problems</i>		273

13. DIELECTRIC PROPERTIES OF MATERIALS **274–291**

13.1	Review of Basic Ideas	274
13.1.1	Electric Dipole	274
13.1.2	Polarizability	274
13.1.3	Polarization Vector	275
13.1.4	Dielectric Constant	275
13.1.5	Dielectric Susceptibility	275
13.2	Dielectric Polarization	276
13.2.1	Electronic Polarization	276
13.2.2	Ionic Polarization	277
13.2.3	Orientation Polarization	278
13.2.4	Space Charge Polarization	278
13.2.5	Total Polarization	279
13.3	Internal Field in a Dielectric	279
13.4	Clausius—Mossotti Relation	281
13.5	Frequency Dependence of Polarization	282
13.6	Dielectric Loss	283
13.7	Ferroelectric Crystals	285
13.8	Piezoelectric Crystals	285
13.8.1	Applications	286

xiv *Contents*

13.9 Applications of Dielectric Materials 286

Solved Examples 287

Review Questions 290

Problems 291

14. MAGNETIC PROPERTIES OF MATERIALS

292–310

14.1 Permeability and Susceptibility 292

14.2 Origin of Magnetic Moment 293

 14.2.1 Orbital Magnetic Moment 293

 14.2.2 Spin Magnetic Moment 294

 14.2.3 Total Magnetic Moment 294

14.3 Classification of Magnetic Materials 295

14.4 Diamagnetism 296

14.5 Paramagnetism 296

14.6 Ferromagnetism 298

 14.6.1 Exchange Interaction 298

 14.6.2 Curie Temperature 299

14.7 Weiss Theory of Ferromagnetism 300

14.8 The Hysteresis Loop 301

14.9 Soft and Hard Magnetic Materials 302

14.10 Antiferromagnetism 303

14.11 Ferrimagnetism 305

 14.11.1 Ferrites 305

14.12 Magnetic Materials for Information Storage 306

Solved Examples 306

Review Questions 309

Problems 310

15. THERMAL PROPERTIES

311–325

15.1 Specific Heat 311

15.2 Einstein Model 312

 15.2.1 High Temperature Limit 313

 15.2.2 Low Temperature Limit 314

15.3 Debye Model 315

15.4 Heat Capacity of Electrons 317

15.5 Lattice Vibrations 318

 15.5.1 One-dimensional Monoatomic Lattice Vibrations 318

15.6 Lattice with Two Atoms Per Primitive Cell 319

15.7 Phonons 320

15.8 Thermal Conductivity 321

Solved Examples 323

Review Questions 325

Problems 325

16. NANOTECHNOLOGY	326–342
16.1 Introduction	326
16.2 Synthesis of Nanomaterials—General Ideas	327
16.3 Laser Induced Evaporation Method	327
16.4 Laser Pulse Method	328
16.5 Metal Nanoclusters – RF Plasma Method	329
16.6 Thermolysis	329
16.7 Chemical Vapour Deposition	330
16.8 Sol–Gel Method	331
16.9 Carbon Clusters	332
16.9.1 Synthesis	332
16.10 Carbon Nanotubes	333
16.10.1 Types of Nanotubes	334
16.11 Fabrication of Carbon Nanotubes	334
16.11.1 Laser Evaporation Method	335
16.11.2 Chemical Vapour Deposition	335
16.12 Nanowires and Nanoribbons	335
16.13 Properties of Nanomaterials	336
16.13.1 Metal Nanoclusters	336
16.13.2 Nanoparticles of Semiconducting Materials	336
16.13.3 Giant Magnetoresistance	337
16.13.4 Carbon Nanotubes	337
16.14 Applications	338
16.15 Medical Applications	341
<i>Review Questions</i>	342
17. ULTRASONICS	343–353
17.1 Ultrasonic Waves	343
17.2 Production of Ultrasonics—Magnetostriction Method	343
17.2.1 Principle	343
17.2.2 Experimental Arrangement	344
17.3 Production of Ultrasonics – Piezoelectric Method	345
17.3.1 Principle	345
17.3.2 Experimental Arrangement	346
17.4 Detection of Ultrasonic Waves	346
17.5 Properties of Ultrasonics	347
17.6 Acoustic Grating	348
17.7 Non-destructive Testing	348
17.7.1 Pulse Echo Method–Flaw Detection	348
17.7.2 Transmission System	349
17.7.3 Thickness Gauging	350
17.7.4 Sonar	350
17.8 Medical Applications of Ultrasonics	350
17.8.1 Ultrasound Imaging	350

xvi *Contents*

17.8.2 Echocardiography	351	
17.8.3 Ultrasound Guided Technique	351	
17.9 Other Applications	351	
<i>Solved Examples</i>	352	
<i>Review Questions</i>	352	
<i>Problems</i>	353	
18. ACOUSTICS OF BUILDINGS		354–364
18.1 Acoustics of Buildings – Basic Requirements	354	
18.2 Sound Intensity	355	
18.3 Absorption Coefficient	355	
18.4 Reverberation	356	
18.4.1 Reverberation Time	356	
18.5 Sabine's Formula for Reverberation Time	357	
18.6 Architectural Acoustics – Factors Affecting and Remedies	358	
18.6.1 Reverberation	358	
18.6.2 Focussing Surfaces	359	
18.6.3 Sufficient Loudness	360	
18.6.4 Absence of Echoes	360	
18.6.5 Resonance	360	
18.6.6 Echelon Effect	360	
<i>Solved Examples</i>	361	
<i>Review Questions</i>	363	
<i>Problems</i>	363	
APPENDICES		365–375
I Abbreviation for Units and Physical Constants	365	
II Fraunhofer Diffraction at a Single Slit	368	
III Metallic Glasses	370	
IV Shape Memory Alloys	372	
V Biomaterials	374	
BIBLIOGRAPHY		377
ANSWERS TO PROBLEMS		379–383
INDEX		385–389



Preface

The book is the resultant of my classroom lectures imparted to the students of various institutions in Kerala including University Institute of Technology, University of Kerala. The book therefore benefits greatly, as it has been shaped on the comments and suggestions provided by the students. As a consequence, the presentation is simple yet complete without sacrificing rigour providing the readers with a sound understanding of different aspects of Engineering Physics. The book is primarily aimed to serve as a textbook for the undergraduate Engineering students of different universities in India and more particularly to students of Jawaharlal Nehru Technical University. However, the arrangement of the different topics are based on the syllabus of the various universities in India.

To begin with, in Chapters 1–5, Oscillation and waves, Electromagnetic theory, Interference, Diffraction and Polarization, the different aspects of wave nature are discussed. A section on anti-reflection coatings is included in interference. Subjects like Crystallography, Superconductivity, Relativity, Quantum mechanics and Statistical mechanics which forms an integral part of the subject are discussed in Chapters 6 to 10. In the chapter on Crystallography, a detailed discussion on cubic lattices, Miller Indices and X-ray diffraction which includes Brag's X-ray spectrometer and X-ray diffraction by powder method are also included. In the chapter on Superconductivity, Meissner effect, Type I and Type II superconductors, BCS theory, Magnetic flux quantization, Josephson junction, high temperature superconductors and various applications are discussed. The concepts and formulation of Quantum mechanics are not elementary in the sense that they are easily understood. In the present book, simple and mathematical techniques have been used to elucidate the physical concepts. A section on quantum mechanical tunneling which is absolutely essential to explain modern electronic devices is discussed in detail.

The exciting and useful areas of Lasers, Fibre optics and Holography are presented in Chapters 11 and 12. Different 3 level and 4 level laser systems which are needed for optical

communications and light pipes are discussed. Chapters 13–15 offer an in-depth presentation of the dielectric, magnetic and thermal properties of materials. The sections on ferro-electric and piezo-electric crystals would be of extreme importance for the study of modern materials. Sections on antiferromagnetism, ferrimagnetism and magnetic materials for information storage are also included. The emerging area of Nanotechnology is discussed in Chapter 16. Different methods for the synthesis of nano-materials and carbon nano tubes along with their properties and applications are dealt with in greater detail. The book closes with the chapters on ultrasonics and acoustic of buildings with sections on non-destructive testing, medical and other applications of ultrasonics.

Learning to solve problems is the basic purpose of a course since it helps in understanding the subject in a better way. Typical worked examples illustrating the concepts involved have been included in almost all chapters. In addition, there are review questions and problems at the end of each chapter. These provide the instructor with enough material for home assignments and classroom discussions. To make things more effective, the quantities are substituted with units. The SI system of units is used throughout with minor exceptions where modifications are found appropriate. The answers to the problems are given at the end of the book.

I am deeply indebted to Prof. Dakshina Murthy, Department of Physics, College of Engineering, JNTU, Kakinada, for going through the entire manuscript and giving extremely valuable suggestions. Also, I would like to express my thanks to my grandchildren Nithin, Cerene, Tina, Zaneta, Juana, Joshua, Tessiya, Lydia and Ezekiel for not stealing away my precious time. Finally, I express my sincere thanks to the publisher PHI Learning, for their unfailing cooperation and for the meticulous processing of the manuscript.

Above all I thank Lord Jesus Christ, who has given me wisdom, power and guidance all through my life.

G. Aruldas

C H A P T E R

1

Oscillations and Waves

The phenomenon of wave motion is prevalent in almost all branches of Physics. Vibrations and wave motions are intimately related topics. Waves have as their source a vibration. In this chapter, we shall discuss certain aspects of vibration followed by the ideas related to travelling waves, wave equation and solution of one and three dimensional problems.

1.1 HARMONIC OSCILLATION—THE BASICS

A motion that is repeated at regular intervals of time is called **periodic motion**. The solutions of the equations of motion of such systems can always be expressed as functions of sines and cosines. Motions described by functions of sines and cosines are often referred to as **harmonic motion**. If the motion is back and forth over the same path, it is called **vibratory or oscillatory**. As oscillation is one round trip of the motion, period T is the time required for one oscillation. The number of oscillations per unit time is the frequency f of oscillation. Hence,

$$T = \frac{1}{f} \quad (1.1)$$

The distance, linear or angular, of the oscillating particle from its equilibrium position is its displacement (linear or angular). The maximum displacement is called the **amplitude** A of the motion.

1.2 SIMPLE HARMONIC MOTION (SHM)

The motion of a particle is said to be **simple harmonic** if it oscillates from an equilibrium position under the influence of a force that is proportional to the distance of the particle from

2 Engineering Physics

the equilibrium position. Also, the force is such that it directs the particle back to its equilibrium position.

1.2.1 Equation of Motion

A familiar example of a particle executing SHM is the motion of a mass m attached to an elastic spring moving on a horizontal frictionless table. When the spring is stretched a distance x from the unextended position, the restoring force F acting on the mass is given by

$$F = -kx \quad (1.2)$$

where k is the proportionality constant, but in the present case [Eq. (1.2)] it is the spring constant k of the spring. The negative sign indicates that the force is directed against the motion, which

is towards the equilibrium position. The acceleration a of the mass is $\left(\frac{d^2x}{dt^2}\right)$. Applying Newton's second law, we get

$$\begin{aligned} -kx &= m \frac{d^2x}{dt^2} \\ m \frac{d^2x}{dt^2} + kx &= 0 \end{aligned} \quad (1.3)$$

Equation (1.3) is the differential equation of motion of a particle executing SHM.

1.2.2 Solution of Equation of Motion

We can rewrite Eq. (1.3) as

$$\frac{d^2x}{dt^2} = -\frac{k}{m}x = -\omega^2x \quad (1.4)$$

where

$$\omega^2 = \frac{k}{m} \quad (1.5)$$

Multiplying both sides of Eq. (1.4) by $2\left(\frac{dx}{dt}\right)$, we get

$$2 \frac{dx}{dt} \frac{d^2x}{dt^2} = -\omega^2 2x \frac{dx}{dt}$$

Integrating with respect to t

$$\left(\frac{dx}{dt}\right)^2 = -\omega^2 x^2 + C$$

where C is a constant.

At the maximum displacement position $x = A$ and velocity $\left(\frac{dx}{dt}\right) = 0$. Hence, $C = \omega^2 A^2$ and

$$\frac{dx}{dt} = \omega \sqrt{(A^2 - x^2)} \quad (1.6)$$

It may be noted from Eqs. (1.4) and (1.6) that in harmonic motion, neither the acceleration nor the velocity of the particle is constant. Rewriting Eq. (1.6) as

$$\frac{dx}{\sqrt{A^2 - x^2}} = \omega dt$$

and integrating

$$\sin^{-1} \frac{x}{A} = \omega t + \phi$$

where ϕ is a constant

$$x = A \sin(\omega t + \phi) \quad (1.7)$$

where A is the amplitude and ϕ is a constant.

First let us have a physical significance of ω . If time t in Eq. (1.7) is increased by $\frac{2\pi}{\omega}$,

$$\begin{aligned} x &= A \sin \left[\omega \left(t + \frac{2\pi}{\omega} \right) + \phi \right] \\ &= A \sin(\omega t + 2\pi + \phi) \\ &= A \sin(\omega t + \phi) \end{aligned}$$

That is, the function simply repeats itself after a time $\frac{2\pi}{\omega}$. Hence, $\frac{2\pi}{\omega}$ is the period of oscillation

T . Since $\omega^2 = \frac{k}{m}$,

$$T = \frac{2\pi}{\omega} = 2\pi \sqrt{\frac{m}{k}} \quad (1.8)$$

The period of a simple harmonic motion is, thus, independent of the amplitude of motion. Frequency

$$f = \frac{1}{T} = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad (1.9)$$

The quantity ω is often called the **angular frequency**. The quantity $(\omega t + \phi)$ is called the **phase** of the motion. The constant ϕ is called the phase constant. If we start counting time when the particle is in its mean position, that is when $x = 0$, $t = 0$ we have from Eq. (1.7) $A \sin \phi = 0$ or $\phi = 0$.

4 Engineering Physics

Hence,

$$x = A \sin \omega t$$

If we start counting time when the particle is in one of its extreme positions, that is when

$$x = A, \quad t = 0 \text{ then,}$$

$$A = A \sin \phi \quad \text{or} \quad \sin \phi = 1 \quad \text{or} \quad \phi = \frac{\pi}{2}$$

Then

$$x = A \sin \left(\omega t + \frac{\pi}{2} \right)$$

If we put $\phi = \phi' + \frac{\pi}{2}$, we get

$$\begin{aligned} x &= A \sin \left(\omega t + \phi' + \frac{\pi}{2} \right) \\ &= A \cos (\omega t + \phi') \end{aligned}$$

Thus, a SHM may be expressed either in terms of a sine or a cosine function. Only the phase constants will have different values in the two cases. What we discussed above is applicable to any body executing simple harmonic motion.

1.2.3 Energy in SHM

Since force is the negative gradient of potential V

$$\therefore -kx = -\frac{dV}{dx} \quad \text{or} \quad dV = kx dx$$

On integration, we get

$$V = \frac{1}{2} kx^2 + C$$

Taking the equilibrium position as the zero of potential energy, $C = 0$, we obtain

$$V = \frac{1}{2} kx^2 \tag{1.10}$$

Total energy E is given by

$$E = \text{kinetic energy} + \text{potential energy}$$

$$E = \frac{1}{2} mv^2 + \frac{1}{2} kx^2$$

At the extreme positions, the velocity is zero and $x = A$. Hence,

$$E = \frac{1}{2} kA^2 \tag{1.10a}$$

At the equilibrium position x will be equal to 0 and the velocity will be maximum, say v_0 . Consequently

$$E = \frac{1}{2}mv_0^2 \quad (1.11)$$

From Eqs. (1.10a) and (1.11) we get

$$v_0 = A\sqrt{\frac{k}{m}} = \omega A \quad (1.11a)$$

1.3 DAMPED HARMONIC OSCILLATION

A harmonic oscillator in which the motion is damped by the action of an additional force, is said to be a **damped harmonic oscillator**. In most of the cases, the damping force is proportional to its velocity, that is the additional frictional force $= b\left(\frac{dx}{dt}\right)$, where the constant b is called the **damping constant**. Including the damping force, the differential equation of a damped harmonic oscillator can be written as

$$\begin{aligned} m\frac{d^2x}{dt^2} &= -kx - b\frac{dx}{dt} \\ m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx &= 0 \end{aligned} \quad (1.12)$$

Replacing $\frac{k}{m}$ by ω^2 and writing $\frac{b}{m} = 2\lambda$, Eq. (1.12) takes the form

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + \omega^2x = 0 \quad (1.13)$$

Writing $\left(\frac{d^2}{dt^2}\right) = D^2$ and $\left(\frac{d}{dt}\right) = D$, the auxiliary equation of Eq. (1.13) is

$$D^2 + 2\lambda D + \omega^2 = 0 \quad (1.14)$$

The roots of this equation are

$$D = \frac{-2\lambda \pm \sqrt{4\lambda^2 - 4\omega^2}}{2} = -\lambda \pm \sqrt{(\lambda^2 - \omega^2)} \quad (1.15)$$

Next, let us examine the nature of the solutions for different cases.

6 Engineering Physics

Case I $\lambda > \omega$: The roots of the auxiliary equation α_1 and α_2 are then real and distinct.

$$\alpha_1 = -\lambda + \sqrt{(\lambda^2 - \omega^2)} \quad ; \quad \alpha_2 = -\lambda - \sqrt{(\lambda^2 - \omega^2)} \quad (1.16)$$

The solution of Eq. (1.13) is then

$$x = c_1 e^{\alpha_1 t} + c_2 e^{\alpha_2 t} \quad (1.17)$$

where c_1 and c_2 are constants and α_1 and α_2 are given by Eq. (1.16). To determine c_1 and c_2 , let the mass is given a displacement x_0 and then released, so that $x = x_0$, $\left(\frac{dx}{dt}\right) = 0$ at $t = 0$. Hence, from Eq. (1.17), we get

$$x_0 = c_1 + c_2 \quad \text{and} \quad c_1 \alpha_1 + c_2 \alpha_2 = 0$$

Simplifying

$$c_1 = -\frac{x_0 \alpha_2}{\alpha_1 - \alpha_2} \quad \text{and} \quad c_2 = \frac{x_0 \alpha_1}{\alpha_1 - \alpha_2} \quad (1.18)$$

Consequently, Eq. (1.17) reduces to

$$x = \frac{x_0}{\alpha_1 - \alpha_2} (\alpha_1 e^{\alpha_2 t} - \alpha_2 e^{\alpha_1 t}) \quad (1.19)$$

which shows that x is always positive and decreases to zero as $t \rightarrow \infty$. The motion is non-oscillatory, and is therefore, referred to as **over damped** or **dead beat** motion.

Case II $\lambda = \omega$: The roots of the auxiliary equation are real and equal, each being $-\lambda$. The general solution of Eq. (1.13) is then

$$x = (c_1 + c_2 t) e^{-\lambda t} \quad (1.20)$$

As in Case I, if $x = x_0$ and $\left(\frac{dx}{dt}\right) = 0$ at $t = 0$ the constants $c_1 = x_0$ and $c_2 = \lambda x_0$.

Hence, the solution of Eq. (1.13) is

$$x = x_0 (1 + \lambda t) e^{-\lambda t} \quad (1.21)$$

Again, x is always positive and decreases to zero as $t \rightarrow \infty$, The nature of motion is similar to that of Case I and it will be clear that Case II separates the non-oscillatory motion of Case I from the interesting oscillatory motion of Case III. Hence, motion in Case II is termed as **critically damped**.

Case III $\lambda < \omega$: When $\lambda < \omega$, the roots of the auxiliary equation are

$$-\lambda \pm i\sqrt{(\omega^2 - \lambda^2)} = -\lambda \pm i\omega' \quad (1.22)$$

where

$$\omega' = \sqrt{\omega^2 - \lambda^2} \quad (1.23)$$

That is, the roots of the auxiliary equation are imaginary and the solution of Eq. (1.13) is

$$x = e^{-\lambda t} (c_1 \cos \omega' t + c_2 \sin \omega' t) \quad (1.24)$$

As in Case I, $x = x_0$, $\frac{dx}{dt} = 0$ at $t = 0$ leads to $c_1 = x_0$ and $c_2 = \frac{\lambda x_0}{\omega'}$. Consequently,

$$x = x_0 e^{-\lambda t} \left(\cos \omega' t + \frac{\lambda}{\omega'} \sin \omega' t \right) \quad (1.25)$$

$$= x_0 e^{-\lambda t} (\cos \omega' t \cos \delta + \sin \omega' t \sin \delta) \quad (1.26)$$

where $\cos \delta = 1$, $\sin \delta = \frac{\lambda}{\omega'}$ or $\tan \delta = \frac{\lambda}{\omega'}$ (1.27)

Now, Eq. (1.26) takes the form

$$x = x_0 e^{-\lambda t} \cos(\omega' t - \delta) \quad (1.28)$$

where $\lambda = \frac{b}{2m}$, $\omega' = \sqrt{\frac{k}{m} - \frac{b^2}{4m^2}}$ and $\delta = \tan^{-1}\left(\frac{\lambda}{\omega'}\right)$ (1.29)

The motion is oscillatory with an amplitude $x_0 \exp(-\lambda t)$ which decreases with $t \rightarrow 0$. From Eq. (1.29) it is obvious that when damping force is present, ω' is less than ω , the angular frequency of the undamped wave. Hence, the period T' is longer when friction is present. Thus, the effect of damping is to increase the period of oscillation and the plot of the displacement x as a function of the time t is shown in Fig. 1.1. Again the theory we discussed is applicable to any body executing a damped harmonic oscillation.

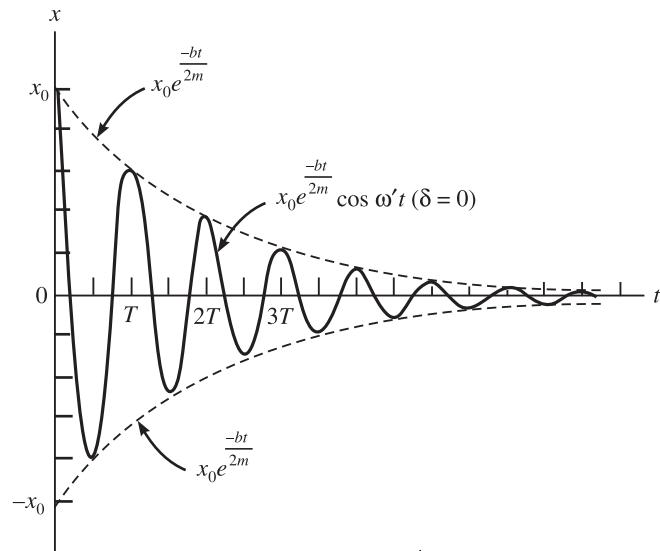


Fig. 1.1 Damped harmonic motion. The amplitude $x_0 e^{-\lambda t}$, where $\lambda = \frac{b}{2m}$ is also plotted.

1.4 FORCED OSCILLATIONS AND RESONANCE

In the previous sections we have been discussing the natural oscillations that occur when a body of mass m is displaced and released. In the absence of damping force the natural frequency of vibration is ω (Eq. 1.5) and when damping is present it is ω' (Eq. 1.23). We shall now investigate the situation when the body is subjected to an oscillatory external force $F_0 \sin \omega''t$. The equation of motion of such a forced oscillator is

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F_0 \sin \omega''t \quad (1.30)$$

$$\frac{d^2x}{dt^2} + 2\lambda \frac{dx}{dt} + \omega^2 x = f_0 \sin \omega''t \quad (1.31)$$

where $\omega = \left(\sqrt{\frac{k}{m}} \right)$ is the natural frequency of the body, and $2\lambda = \left(\frac{b}{m} \right)$ is the damping constant for unit mass and $f_0 = \frac{F_0}{m}$.

When external periodic force is applied, both the damping and forced terms contribute to the motion of the oscillator and a tussle occurs between the contributions of both. After some initial erratic situations, ultimately, the system reaches a steady state. The oscillations that result are called the **forced oscillations**. These forced oscillations have the frequency of the external force and not that of the natural frequency of the body. Let us consider a solution of the type when the steady state is reached,

$$x = A \sin(\omega''t - \theta) \quad (1.32)$$

for the equation of motion, Eq. (1.31). Then

$$\frac{dx}{dt} = A\omega'' \cos(\omega''t - \theta)$$

$$\frac{d^2x}{dt^2} = -A\omega''^2 \sin(\omega''t - \theta)$$

Substituting these values in Eq. (1.31), we get

$$-A\omega''^2 \sin(\omega''t - \theta) + 2\lambda A\omega'' \cos(\omega''t - \theta) + \omega^2 A \sin(\omega''t - \theta) = f_0 \sin[(\omega''t - \theta) + \theta]$$

where we have added and subtracted θ in the right hand side term. Expanding the right side, we obtain

$$\begin{aligned} & -A\omega''^2 \sin(\omega''t - \theta) + 2\lambda A\omega'' \cos(\omega''t - \theta) + \omega^2 A \sin(\omega''t - \theta) \\ &= f_0 \sin(\omega''t - \theta) \cos \theta + f_0 \cos(\omega''t - \theta) \sin \theta \\ & (-A\omega''^2 - f_0 \cos \theta + \omega^2 A) \sin(\omega''t - \theta) + (2\lambda A\omega'' - f_0 \sin \theta) \cos(\omega''t - \theta) = 0 \end{aligned} \quad (1.33)$$

For this equation to hold good for all values of t , the coefficients of the terms $\sin(\omega''t - \theta)$ and $\cos(\omega''t - \theta)$ must vanish separately.

$$\text{i.e. } -A\omega''^2 - f_0 \cos \theta + \omega^2 A = 0 \quad \text{and} \quad 2\lambda A\omega'' - f_0 \sin \theta = 0 \quad (1.34)$$

$$-A\omega''^2 + \omega^2 A = f_0 \cos \theta \quad \text{and} \quad 2\lambda A\omega'' = f_0 \sin \theta \quad (1.35)$$

Squaring two Eqs. (1.34) and (1.35) and adding, we get

$$A^2[(-\omega''^2 + \omega^2)^2 + 4\lambda^2\omega''^2] = f_0^2$$

$$A = \frac{f_0}{\sqrt{(\omega^2 - \omega''^2)^2 + 4\lambda^2\omega''^2}} \quad (1.36)$$

which gives the amplitude of the forced oscillation. From Eq. (1.35), we obtain

$$\tan \theta = \frac{2\lambda\omega''}{\omega^2 - \omega''^2} \quad (1.37)$$

This gives the phase difference between the forced oscillator and the applied force. Substituting the value of A in Eq. (1.32), we obtain

$$x = \frac{f_0}{\sqrt{(\omega^2 - \omega''^2)^2 + 4\lambda^2\omega''^2}} \sin(\omega''t - \theta) \quad (1.38)$$

That is, the system is undamped and vibrate with the frequency of the driving force ω'' , but with a phase lag of θ given by Eq. (1.37).

The denominator of the expression for amplitude A is large when the frequency of the driving force ω'' is very different from the natural frequency ω of the undamped system. That means that the amplitude of the resulting motion is small. However, as the driving frequency ω'' approaches the natural frequency ω , the amplitude increases and reaches the maximum value when both are nearly equal. The phenomenon is called **resonance** and the frequency ω'' which gives the maximum amplitude is called the **resonant frequency**. The amplitude of the forced vibration depends on the damping force also. The greater the damping force, larger is the denominator in Eq. (1.43), resulting in a smaller amplitude. Figure 1.2 shows five curves giving the amplitude of the forced vibration as a function of $\frac{\omega''}{\omega}$. Each one corresponds to a different value of damping constant λ .

The solutions of forced vibrations is useful in acoustic systems and alternating current circuits as well as in mechanics.

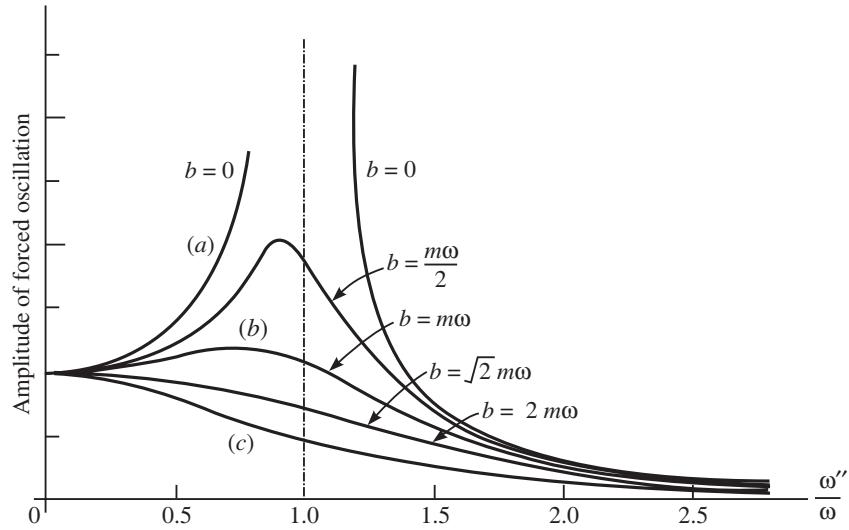


Fig. 1.2 Amplitude of a forced harmonic oscillator versus $\frac{\omega''}{\omega}$ for five values of b .

In Fig. 1.2, curve (a) is for zero damping and, curve (b) is for intermediate damping and curve (c) is for high damping.

1.5 QUALITY FACTOR

Quality factor decides the quality of the oscillator as far as damping is concerned. For the damped harmonic oscillator, the displacement x is given by Eq. (1.28). The velocity of the particle at a particular instant

$$\frac{dx}{dt} = -x_0 e^{-\lambda t} [\lambda \cos(\omega't - \delta) + \omega' \sin(\omega't - \delta)]$$

Kinetic energy (K.E.) of the oscillator is given by

$$\begin{aligned} \text{K.E.} &= \frac{1}{2} m \left(\frac{dx}{dt} \right)^2 \\ &= \frac{1}{2} m x_0^2 e^{-2\lambda t} [\lambda \cos(\omega't - \delta) + \omega' \sin(\omega't - \delta)]^2 \end{aligned} \quad (1.39)$$

When expanded, the expression will have a term in $\cos^2(\omega't - \delta)$ another term in $\sin^2(\omega't - \delta)$ and a cross term. The average value of the terms $\cos^2(\omega't - \delta)$ and $\sin^2(\omega't - \delta)$ over a period $= \frac{1}{2}$. The contribution from the cross term over a period is zero. Hence, average kinetic energy of the particle over one cycle at the given instant t is

$$\text{K.E.} = \frac{1}{2} mx_0^2 e^{-2\lambda t} \left(\frac{\lambda^2}{2} + \frac{\omega'^2}{2} \right) \quad (1.40)$$

$$\approx \frac{1}{4} mx_0^2 e^{-2\lambda t} \omega'^2 \quad (1.41)$$

since $\lambda^2 \ll \omega'^2$. Using Eqs. (1.10) and (1.5), the potential energy (P.E.) of the oscillating particle at a given time t is

$$\text{P.E.} = \frac{1}{2} m \omega'^2 x_0^2 e^{-2\lambda t} \cos^2(\omega' t - \delta) \quad (1.42)$$

$$\text{Average P.E. over a time period} = \frac{1}{4} m \omega'^2 x_0^2 e^{-2\lambda t} \quad (1.43)$$

Total energy

$$E = \text{K.E.} + \text{P.E.} = \frac{1}{2} m \omega'^2 x_0^2 e^{-2\lambda t} \quad (1.44)$$

Power dissipation P is rate of loss of total energy with time

$$\begin{aligned} P &= -\frac{dE}{dt} = mx_0^2 \omega'^2 \lambda e^{-2\lambda t} \\ &= 2\lambda E \end{aligned} \quad (1.45)$$

Quality factor Q is defined by

$$\begin{aligned} Q &= \frac{2\pi \times \text{Energy of the oscillator}}{\text{Energy lost per cycle}} \\ &= \frac{2\pi E}{\left(\frac{\text{Energy lost}}{s} \right) \times T} = \frac{2\pi E}{2\lambda ET} = \frac{\pi}{\lambda T} \end{aligned}$$

$$\text{Since } T = \frac{2\pi}{\omega'} \quad (1.45a)$$

$$Q = \frac{\pi \omega'}{\lambda 2\pi} = \frac{\omega'}{2\lambda} \quad (1.45a)$$

The quality factor Q will be large if the damping coefficient λ is small and Q will be small if λ is large. In other words, quality factor represents the efficiency of the oscillator.

1.6 TYPES OF WAVES

When we throw a stone in a pool of water, waves form and travel outward. When we shout, sound waves are generated and travel in all directions. Light waves are generated and travel outward when an electric bulb is switched on. In all these cases, there is a transfer of energy in the form of waves. Different types of waves exist in the different branches of Physics. We shall discuss some of these in the next sections.

1.6.1 Mechanical Waves

Mechanical waves originate from the vibration of some portion of an elastic medium about its equilibrium position. Because of the elastic properties of the medium, the disturbance is transmitted from one layer to the another. This disturbance or wave progresses through the medium. An important point is that the medium itself does not move as a whole, only the various parts of the medium vibrate. These mechanical waves are characterized by the transport of energy through matter. Hence, a medium having elasticity is necessary for the transmission of mechanical waves. Sound waves, waves on water surface, seismic waves etc. are some of the familiar examples.

1.6.2 Electromagnetic Waves

Electromagnetic waves are always of fields, not of matter as waves on water or a rope. An electromagnetic field is characterized by its electric field strength **E** and magnetic field intensity **B**. Electromagnetic waves are produced by oscillating electric charges. The electric and magnetic fields associated with the wave at any point are perpendicular to each other and to the direction of propagation. The field strengths vary from a maximum in one direction, to zero, to a maximum in the other direction. The electric and magnetic fields are in phase, that is, each of them is zero at the same points in space. Light waves, X-rays, infrared rays, microwaves, radio and TV waves come under this category. We will be learning more about electromagnetic waves in Chapter 2.

1.6.3 Other Types of Waves

Besides these, there can also be other types of waves. When an earthquake occurs, earthquake waves consisting of both transverse and longitudinal waves are produced. The transverse waves travel through the body of the earth. Scientists detected longitudinal waves on the surface of the earth as well as diametrically across the earth. From the first observation, it is inferred that the earth's outer core is molten. The second observation implies that the core of the earth must be a liquid. Another type is the surface waves that travel along the boundary between two materials. A wave on water is a surface wave that moves on the boundary between air and water. The motion of each particle on the surface of water is circular or elliptical, so it is a combination of longitudinal and transverse motions. Well below the surface, the motion is only longitudinal. Surface waves are also set up during earthquakes. The waves that travel along the surface are responsible for the damage.

1.7 GENERAL FORM OF A TRAVELLING WAVE

Consider a transverse wave pulse travelling in a string stretched in the x -direction. Let $u(x, t)$ be the transverse displacement of the string at x . At any instant of time, say $t = 0$, the displacement u is a function of x . That is

$$u(x, 0) = F(x), \quad t = 0 \quad (1.46)$$

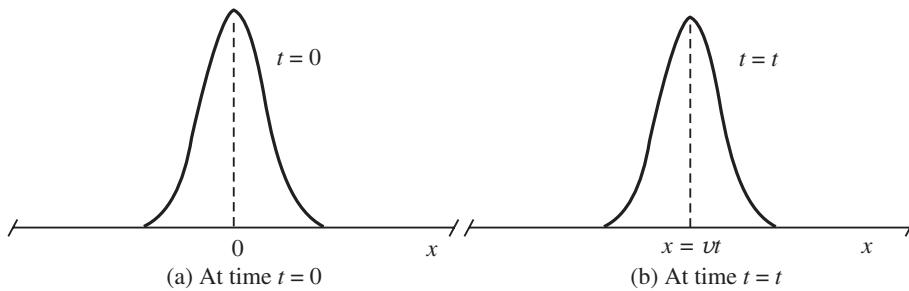


Fig. 1.3 The waveform (pulse) on the string.

Figure 1.3(a) shows the waveform (pulse) on the string. At time t , the wave has travelled a distance vt to the right, where v is the wave velocity of the pulse which is assumed to be constant. The equation that describes the wave pulse at time t is then given by

$$u(x, t) = F(x - vt), \quad t = t \quad (1.47)$$

Figure 1.3(b) shows the waveform on the string at time t . One of the characteristics of a wave is its propagation without change in shape. That is, the form of the pulse has identical shapes at $t = 0$ and at $t = t$. For $F(x - vt)$ to have the same form, $(x - vt)$ must have a constant value. To keep $(x - vt)$ constant, as t increases x must increase accordingly. Hence, $F(x - vt)$ represents a wave travelling in the positive x -direction. To get the shape of the waveform, the exact form of the function is needed. In the same way, for a wave travelling in the negative x -direction, we have

$$u = F(x + vt) \quad (1.48)$$

Here, the position x of some fixed phase $(x + vt)$ of the wave decreases as time increases. For the wave travelling to the right, for a particular phase, we get

$$x - vt = \text{Constant} \quad \text{or} \quad \frac{dx}{dt} = v \quad (1.49)$$

In the same way, for a wave travelling to the left, $-v$ is its phase velocity. The general solution is the sum of the two waves. The discussion just presented holds for longitudinal waves also.

$$u = F_1(x + vt) + F_2(x - vt) \quad (1.50)$$

where F_1 is a function of $(x + vt)$ and F_2 is a function of $(x - vt)$

1.8 HARMONIC WAVE

We now consider the propagation of a periodic wave in which the displacement is a function of x and t . At $t = 0$, let it be of the type

$$u = A \sin \frac{2\pi x}{\lambda} \quad (1.51)$$

14 Engineering Physics

Its shape is shown in Fig. 1.4. The maximum displacement A is the amplitude of the wave. The quantity $\frac{2\pi x}{\lambda}$ is known as the phase of the wave. The value of the transverse displacement at $x + \lambda$ is given by

$$\begin{aligned} u &= A \sin \frac{2\pi(x + \lambda)}{\lambda} = A \sin \left(\frac{2\pi x}{\lambda} + 2\pi \right) \\ &= A \sin \frac{2\pi x}{\lambda} \end{aligned}$$

Hence, the value of transverse displacement will be the same at $x, x + \lambda, x + 2\lambda$ and so on. The quantity λ is called the **wavelength of the wave train**. Wavelength represents the distance between two adjacent points in the wave having the same phase. Let the wave travels with the velocity v in the positive x -direction. Then the equation that represents the wave at the time t is

$$u = A \sin \frac{2\pi}{\lambda} (x - vt) \quad (1.52)$$

It may be noted that it has the form required for a travelling wave, Eq. (1.47). The quantity $2\pi \frac{(x - vt)}{\lambda}$ is the phase at time t . The time required for the wave to travel a distance of one wavelength λ is the period T :

$$\frac{\lambda}{v} = T \quad \text{or} \quad \lambda = vT \quad (1.53)$$

The frequency v and period T are related by the expression

$$v = \frac{1}{T} \quad (1.54)$$

Replacing v in Eq. (1.52) using Eq. (1.53)

$$u = A \sin 2\pi \left(\frac{x}{\lambda} - \frac{t}{T} \right) \quad (1.55)$$

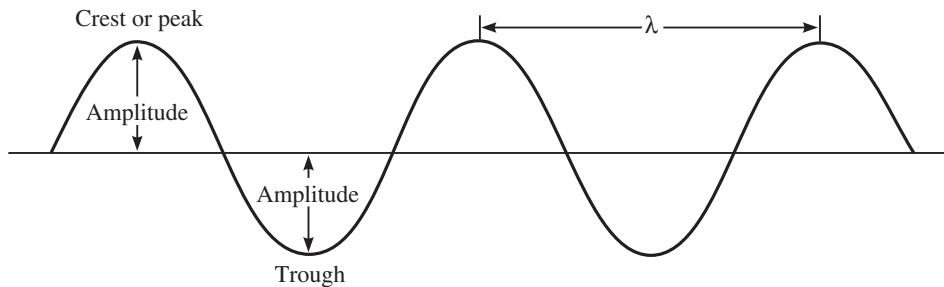


Fig. 1.4 A sine wave travelling in the positive x -direction.

In Fig. 1.4, the quantity A is the amplitude, and λ is the wavelength of the wave.

This form of the equation of the wave immediately tells us that u has the same value at $x, x + \lambda, x + 2\lambda, \dots$, at any given time and that u has the same value at times $t, t + T$,

$t + 2T, \dots$, at any given position. The **wave number** k and the **angular frequency** ω are defined by the relations

$$k = \frac{2\pi}{\lambda} \quad \text{and} \quad \omega = \frac{2\pi}{T} = 2\pi\nu \quad (1.56)$$

In terms of these quantities, Eq. (1.55) reduces to

$$u = A \sin(kx - \omega t) \quad (1.57)$$

The expression for the wave velocity v , Eq. (1.53), now becomes

$$v = \frac{\lambda}{T} = \frac{\omega}{k} \quad (1.58)$$

which is obviously a constant. For a sine wave travelling in the negative x -direction.

$$u = A \sin(kx + \omega t) \quad (1.59)$$

In deriving Eq. (1.57) we have assumed that the displacement u is zero at $x = 0$ and at time $t = 0$. This need not be the case. The general expression for a sinusoidal wave travelling to the right can be written as

$$u = A \sin(kx - \omega t + \phi) \quad (1.60)$$

1.9 WAVE VELOCITY AND PARTICLE VELOCITY

When a harmonic wave travels in a medium, the particles of the medium undergo simple harmonic oscillations. Hence, the velocity of the particles of the medium vary with time whereas the wave velocity, Eq. (1.58), is a constant. The particle velocity v_p is defined by

$$v_p = \frac{du}{dt} \quad (1.61)$$

We shall now find the difference between the wave velocity v and the particle velocity v_p . For any wave travelling in the x -direction with a particular phase of the wave, $(x - vt) = \text{constant}$. On differentiating, we get

$$dx - vdt = 0 \quad \text{or} \quad v = \frac{dx}{dt} \quad (1.62)$$

As mentioned, the wave velocity v is really the phase velocity of the wave. To fix our attention on a particle at a particular position, x has to be constant in Eq. (1.57).

$$\begin{aligned} v_p &= \frac{du}{dt} = -\omega A \cos(kx - \omega t) \\ \frac{du}{dx} &= kA \cos(kx - \omega t) \end{aligned} \quad (1.63)$$

Hence

$$\begin{aligned} v_p &= -\omega \frac{1}{k} \frac{du}{dx} = -2 D \frac{1}{2} \frac{du}{dx} \\ &= -D\lambda \frac{du}{dx} = -v \frac{du}{dx} \end{aligned}$$

16 Engineering Physics

Hence the particle velocity at any instant $v_p = -$ wave velocity x slope of the displacement curve at that instant.

Evidently, as expected the particle velocity v_p changes with time. The acceleration a of the particle at x is

$$a = \frac{dv_p}{dt} = \frac{d^2u}{dt^2} = -\omega^2 A \sin(kx - \omega t) = -\omega^2 u \quad (1.64)$$

This also shows that each particle executes a simple harmonic motion.

1.10 ENERGY TRANSPORT IN WAVE MOTION

The source of energy is the vibrating object which sends out the waves. As wave travels through a medium, energy derived from the source is transferred as vibrational energy from one particle to another of the medium resulting a transport of both kinetic energy of motion of particles and potential energy of deformation of the medium.

Consider a wave propagating in a medium in the x -direction. The time variation of the displacement of a particle of the medium is given by Eq. (1.57). The instantaneous velocity of the particle v_p is given by Eq. (1.63). The kinetic energy T is then given by

$$T = \frac{1}{2} m v_p^2 = \frac{1}{2} m \omega^2 A^2 \cos^2(kx - \omega t) \quad (1.65)$$

As particles execute SHM, total energy E must equal the maximum kinetic energy:

$$E = T_{\max} = \frac{1}{2} m \omega^2 A^2 \quad (1.66)$$

$$= 2\pi^2 m v^2 A^2 \quad (1.67)$$

The energy transported by a wave is proportional to the square of the amplitude, and it does not change with time.

The mass $m = \rho V$ where ρ is the density of the medium and V is the volume of a small slice of the medium. If a is the cross sectional area through which the wave travels, $V = al$ where l is the distance, the wave travels in time t . As $l = vt$

$$m = \rho V = \rho a l = \rho a v t \quad (1.68)$$

With this value of m , Eq. (1.67) reduces to

$$E = 2\pi^2 \rho a v t v^2 A^2 \quad (1.69)$$

The average rate of energy transported is power P

$$P = \frac{E}{t} = 2\pi^2 \rho a v v^2 A^2 \quad (1.70)$$

The intensity I of a wave is defined as power transported across unit area perpendicular to the direction of energy flow given as

$$I = \frac{P}{a} = 2\pi^2 \rho v v^2 A^2 \quad (1.71)$$

This expression is valid for all types of waves including the lightwave or sound wave in space.

1.11 THE GENERAL WAVE EQUATION

1.11.1 One-dimensional Wave Equation

The displacement u of any point on a sinusoidal wave in both space and time is given by Eq. (1.52)

$$u = A \sin \frac{2\pi}{\lambda} (x - vt) \quad (1.72)$$

We shall now derive one dimensional wave equation that connects variations in time and space to the velocity of propagation of the wave.

Differentiating Eq. (1.72) with respect to x and t we get

$$\frac{\partial u}{\partial x} = \frac{2\pi}{\lambda} A \cos \frac{2\pi}{\lambda} (x - vt) \quad (1.73)$$

$$\frac{\partial u}{\partial t} = -\frac{2\pi v}{\lambda} A \cos \frac{2\pi}{\lambda} (x - vt) \quad (1.74)$$

Combining both the equations and dropping the common factors, we get

$$\frac{\partial u}{\partial x} = -\frac{1}{v} \frac{\partial u}{\partial t} \quad (1.75)$$

which is the differential equation describing the wave. If the wave were to move in the negative x -direction.

$$u = A \sin \frac{2\pi}{\lambda} (x + vt) \quad (1.76)$$

Consequently, the differential equation describing a wave travelling to the left will be

$$\frac{\partial u}{\partial x} = \frac{1}{v} \frac{\partial u}{\partial t} \quad (1.77)$$

If we take a second derivative equation as detailed below, it would hold good for any harmonic wave independent of the direction of travel. Taking the second derivative of Eq. (1.75)

$$\frac{\partial}{\partial x} \frac{\partial u}{\partial x} = -\frac{1}{v} \frac{\partial}{\partial x} \frac{\partial u}{\partial t} \quad (1.78)$$

$$= -\frac{1}{v} \frac{\partial}{\partial t} \frac{\partial u}{\partial x} \quad (1.79)$$

where we have changed the order of differentiation on the right side. Substituting the value of $\frac{\partial u}{\partial x}$, we get

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} \quad (1.80)$$

which is the required differential equation.

1.11.2 Three-dimensional Wave Equation

To generalise Eq. (1.80) to waves propagating in any direction, we have to extend the left hand side term in Eq. (1.80) to the y and z axes. Then, the differential Eq. (1.80) becomes

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} \quad (1.81)$$

$$\nabla^2 u = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} \quad (1.82)$$

where the Laplacian operator ∇^2 is defined by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (1.83)$$

Thus, Eq. (1.81) gives the differential equation for a wave propagating in any direction in three dimensional space. The quantity $u(x, y, z, t)$ could be a mechanical displacement or the field components of an electromagnetic wave or the potential of an electrical transmission line.

1.12 SOLUTION OF THE ONE-DIMENSIONAL WAVE EQUATION

The one-dimensional equation is Eq. (1.80) where v is the wave velocity in the medium. In order to find the solution of Eq. (1.80), we try a solution of the form

$$u(x, t) = X(x) T(t) \quad (1.84)$$

where $X(x)$ is a function of x and $T(t)$ is a function of t . Substituting the assumed solution in Eq. (1.80), we get

$$T \frac{d^2 X}{dx^2} = \frac{X}{v^2} \frac{d^2 T}{dt^2}$$

Dividing the Eq. throughout by XT , we get

$$\frac{1}{X} \frac{d^2 X}{dx^2} = \frac{1}{v^2} \frac{1}{T} \frac{d^2 T}{dt^2} \quad (1.85)$$

A change in x will not change the right side of the equation and a change in t will not change the left side. Hence, each side must be equal to a constant, $-k^2$. Consequently,

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -k^2 \quad \text{or} \quad \frac{d^2 X}{dx^2} = -k^2 X \quad (1.86)$$

$$\frac{1}{v^2} \frac{1}{T} \frac{d^2 T}{dt^2} = -k^2 \quad \text{or} \quad \frac{d^2 T}{dt^2} = -k^2 v^2 T \quad (1.87)$$

Writing $k^2 v^2 = \omega^2$, Eq. (1.87) becomes

$$\frac{d^2 T}{dt^2} = -\omega^2 T \quad (1.88)$$

Equations (1.86) and (1.88) are standard linear differential equations with constant coefficients. The solutions can be written in terms of sine and cosine functions or in terms of exponential functions. The solution of Eqs. (1.86) and (1.88) are, gives as

$$X(x) = \text{Constant } e^{\pm ikx} \quad (1.89)$$

$$T(t) = \text{Constant } e^{\pm i\omega t} \quad (1.90)$$

From Eqs. (1.84), (1.89) and (1.90), we can write the solution of the one dimensional wave equation as

$$u(x, t) = A \exp[i(kx \pm \omega t)] \quad (1.91)$$

where A is a constant to be determined by the initial condition. Equation (1.91) is of the same form as Eq. (1.55). Also, it is evident that the constant ω has the meaning of angular frequency and k is the wave vector.

1.13 SOLUTION OF THE THREE-DIMENSIONAL WAVE EQUATION

The three-dimensional wave equation Eq. (1.81) can also be solved by the method of separation of variables. For that, we write,

$$u(x, y, z, t) = X(x) Y(y) Z(z) T(t) \quad (1.92)$$

where $X(x)$, $Y(y)$, $Z(z)$ and $T(t)$ are functions of x , y , z and t , respectively. Substituting Eq. (1.92) in Eq. (1.81), we get

$$YZT \frac{d^2 X}{dx^2} + XZT \frac{d^2 Y}{dy^2} + XYT \frac{d^2 Z}{dz^2} = \frac{1}{v^2} XYZ \frac{d^2 T}{dt^2} \quad (1.93)$$

Dividing Eq. (1.93) throughout by $XYZT$, we get

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = \frac{1}{v^2} \frac{1}{T} \frac{d^2 T}{dt^2} \quad (1.94)$$

The right hand side of Eq. (1.94) is a function of t alone and the left hand side is a function of x , y and z . This is possible only when each side is a constant, say $-k^2$. Then, we get

$$\frac{1}{v^2} \frac{1}{T} \frac{d^2 T}{dt^2} = -k^2 \quad (1.95)$$

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = -k^2 \quad (1.96)$$

20 Engineering Physics

Equation (1.96) can also be written as

$$\frac{1}{X} \frac{d^2X}{dx^2} = -k^2 - \frac{1}{Y} \frac{d^2Y}{dy^2} - \frac{1}{Z} \frac{d^2Z}{dz^2} \quad (1.97)$$

Left hand side of Eq. (1.97) is a function of x alone, and the right hand side is a function of y and z which is possible only when each side is a constant, say $-k_x^2$. Equation (1.97) reduces into two equations as follows:

$$\frac{1}{X} \frac{d^2X}{dx^2} = -k_x^2 \quad (1.98)$$

$$-k^2 - \frac{1}{Y} \frac{d^2Y}{dy^2} - \frac{1}{Z} \frac{d^2Z}{dz^2} = -k_x^2 \quad (1.99)$$

Following the same procedure, Eq. (1.99) reduces to

$$\frac{1}{Y} \frac{d^2Y}{dy^2} = -k_y^2 \quad (1.100)$$

$$\text{and} \quad \frac{1}{Z} \frac{d^2Z}{dz^2} = -k_z^2 \quad (1.101)$$

where k_x^2, k_y^2, k_z^2 and k^2 are constants related by the equation

$$k^2 = k_x^2 + k_y^2 + k_z^2 \quad (1.102)$$

Equation (1.81), thus reduces to the four one dimensional equations given by Eqs. (1.95), (1.98), (1.100) and (1.101).

The solutions of these four equations can be written in terms of sine and cosine functions or in terms of exponential functions. Writing $k^2 v^2 = \omega^2$ and rearranging, Eq. (1.95), we get

$$\frac{d^2T}{dt^2} = -k^2 v^2 \quad T = -\omega^2 T \quad (1.103)$$

Since k, v and ω are constants the solution of Eq. (1.103) becomes

$$T = \text{constant } e^{\pm i\omega t} \quad (1.104)$$

Since v is the wave velocity given by $v = \lambda\nu$, the constant ω has the meaning of an angular frequency. In the same manner we can write the solutions for Eqs. (1.98), (1.100) and (1.101). Both negative and positive exponential solutions are allowed. We have $k^2 = k_x^2 + k_y^2 + k_z^2$. If k_x, k_y and k_z are taken to be real, they may be considered as the components of a vector \mathbf{k} . In such a case $k_x x + k_y y + k_z z = \mathbf{k} \cdot \mathbf{r}$ where $\mathbf{r} = \hat{i}x + \hat{j}y + \hat{k}z$. Consequently, we can write

$$X(x) Y(y) Z(z) = \text{Constant } e^{i\mathbf{k} \cdot \mathbf{r}} \quad (1.105)$$

Substituting Eqs. (1.104) and (1.105) in Eq. (1.92), we get the solution of the three dimensional wave equation as

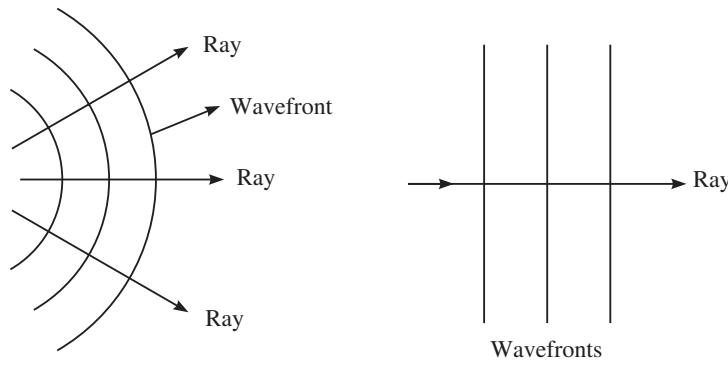
$$u(x, y, z, t) = A \exp [i(\mathbf{k} \cdot \mathbf{r} \pm \omega t + \phi)] \quad (1.106)$$

where A and ϕ are constants to be determined by the initial conditions.

1.14 PLANE AND SPHERICAL WAVES

Consider a three dimensional wave diverging from a point source. All the points on the surface of a sphere with the point source as the centre have the same amplitude and same phase. The wavefronts will be spherical. If the medium is homogenous and isotropic, the direction of propagation is always at right angles to the wavefront. Such waves are called **spherical waves**. Near the point source, the wavefronts are spheres and the rays are radial lines leaving the source in all directions. Figure 1.5(a) shows the projection of the spherical wavefronts near the source. Mathematically, a diverging spherical wave is represented by

$$\Psi = \frac{A}{r} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (1.107)$$



(a) Spherical wavefronts
(only projection shown) near a source (b) Plane wavefronts far from the source

Fig. 1.5 Projection of spherical and plane wavefronts.

The factor $\frac{1}{r}$ in the expression which implies that the amplitude of a spherical wave decreases inversely as r . Hence, the intensity will fall off as $\frac{1}{r^2}$.

Far away from the source, the wavefronts have lost almost all their curvature and are nearly plane, Fig. 1.5(b). Such waves are called **plane waves**. Mathematically, the plane wave travelling in some arbitrary direction is represented by

$$\Psi = Ae^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (1.108)$$

1.15 TRANSVERSE VIBRATIONS OF A STRETCHED STRING

The study of the vibrations of a stretched string is an important problem in the theory of wave motion. Though the differential equation for the transverse vibration of a stretched string can be written from Eq. (1.80), to get better insight we prefer to derive it from fundamentals.

Consider a perfectly flexible uniform string stretched between two points by a constant tension T . Let m be the mass per unit length of the wire. Let us assume that the undisturbed position of the string is the x -axis and that the motion is confined to the xy plane. Consider the motion of an element AB of length dx (undisturbed) as shown in Fig. 1.6.

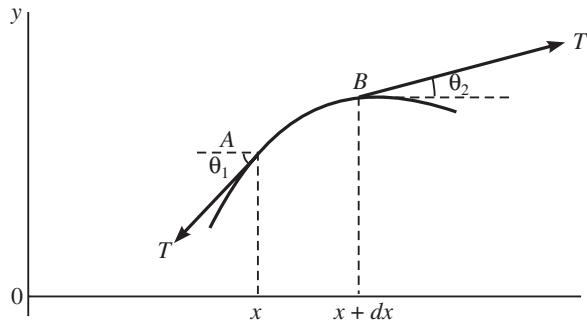


Fig. 1.6 Transverse vibration of a stretched string.

θ_1 and θ_2 are the angles the tension T make with A and B at the x -axis. The net force in the y -direction F_y is given by

$$F_y = T \sin \theta_2 - T \sin \theta_1 \quad (1.109)$$

For small oscillations,

$$\begin{aligned} F_y &= T \tan \theta_2 - T \tan \theta_1 \\ F_y &= T \left(\frac{\partial y}{\partial x} \right)_{x+dx} - T \left(\frac{\partial y}{\partial x} \right)_x \end{aligned} \quad (1.110)$$

Using Taylor's theorem and neglecting higher order terms, we get

$$\left(\frac{\partial y}{\partial x} \right)_{x+dx} = \left(\frac{\partial y}{\partial x} \right)_x + \frac{\partial}{\partial x} \left(\frac{\partial y}{\partial x} \right)_x dx \quad (1.111)$$

Substituting Eq. (1.111) in Eq. (1.110), we get

$$F_y = T \frac{\partial}{\partial x} \left(\frac{\partial y}{\partial x} \right)_x dx = T \frac{\partial^2 y}{\partial x^2} dx$$

Applying Newton's second law of motion, we get

$$\begin{aligned} T \frac{\partial^2 y}{\partial x^2} dx &= m dx \frac{\partial^2 y}{\partial t^2} \\ \frac{\partial^2 y}{\partial x^2} &= \frac{m}{T} \frac{\partial^2 y}{\partial t^2} = \frac{1}{\frac{T}{m}} \frac{\partial^2 y}{\partial t^2} \end{aligned} \quad (1.112)$$

Comparing Eq. (1.112) with the general wave equation, Eq. (1.80) we can write

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 y}{\partial t^2} \quad (1.113)$$

where

$$v = \sqrt{\frac{T}{m}} \quad (1.114)$$

Equation (1.113) is the differential equation for the vibrations of a stretched string. It is easy to show that v has the dimensions of length/time. Hence, we can identify v as the velocity of the wave in the string.

SOLVED EXAMPLES

Example 1.1 A particle describes SHM in a line 4 cm long. Its velocity when passing through the centre of the line is 12 cm/s. Find the period of motion.

Solution: Amplitude of motion $A = \frac{4\text{cm}}{2} = 2\text{cm}$

From Eq. (1.6) velocity v at the mean position $x = 0$ is

$$v = \omega A \quad \text{or} \quad \omega = \frac{12\text{cm/s}}{2\text{cm}} = 6\text{s}^{-1}$$

Time period

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{6\text{s}^{-1}} = \frac{\pi}{3}\text{s}$$

Example 1.2 A particle executing SHM has a period of 0.1 s and amplitude 4 cm. Find its acceleration when it is 0.2 cm from its mean position and its maximum velocity.

Solution: Period

$$T = \frac{2\pi}{\omega}$$

$$\omega = \frac{2\pi}{0.1s} = 20\pi\text{s}^{-1}$$

From Eq. (1.4)

$$\begin{aligned} \text{Acceleration} &= \omega^2 x = (20\pi\text{s}^{-1})^2 \cdot 0.2\text{cm} \\ &= 790.3 \text{ cm s}^{-2} \end{aligned}$$

24 Engineering Physics

Maximum velocity is when the particle is in its mean position

$$\begin{aligned}\text{Maximum velocity} &= \omega A = (20\pi s^{-1}) 4 \text{ cm} \\ &= 251.2 \text{ cm s}^{-1}\end{aligned}$$

Example 1.3 A damped vibrating system, starting from rest, reaches the first amplitude of 40 cm which reduces to 4 cm in that direction after 100 oscillations. If the period of each oscillation is 2.5 s, find the damping constant.

Solution: The amplitude of vibration $= x_0 e^{-\lambda t}$

From the mean position, the time taken to reach the first amplitude is $\frac{T}{4}$. Hence

$$\begin{aligned}\text{Amplitude} &A_1 = x_0 e^{-\lambda \frac{T}{4}} \\ &A_2 = x_0 e^{-\lambda \left(\frac{T}{4} + T\right)} \\ &A_{n+1} = x_0 e^{-\lambda \left(\frac{T}{4} + nT\right)} \\ &\frac{A_1}{A_{n+1}} = e^{n\lambda T}\end{aligned}$$

$$\frac{40 \text{ cm}}{4 \text{ cm}} = e^{100\lambda T} \quad \text{or} \quad \ln 10 = 100 \lambda T$$

$$\lambda = \frac{\ln 10}{100 \times 2.5} = 0.92 \times 10^{-2}$$

Damping constant $\lambda = 0.92 \times 10^{-2}$

Example 1.4 A particle executing SHM along a straight line has velocity 16 cm/s and 12 cm/s when passing through points 3 cm and 4 cm from the mean position, respectively. Find (i) the amplitude and (ii) period of oscillation.

Solution: From Eq. (1.6), we get

$$v = \omega \sqrt{A^2 - x^2} \quad (\text{i})$$

$$16 \text{ cm s}^{-1} = \omega \sqrt{A^2 - 9} \quad (\text{ii})$$

$$12 \text{ cm s}^{-1} = \omega \sqrt{A^2 - 16} \quad (\text{iii})$$

Dividing Eq. (ii) by Eq. (iii), we get

$$\frac{4}{3} = \frac{\sqrt{A^2 - 9}}{\sqrt{A^2 - 16}}$$

$$\frac{A^2 - 9}{A^2 - 16} = \frac{16}{9} \quad \text{or} \quad 7A^2 = 175$$

$$A = 5 \text{ cm}$$

From Eq. (ii), we get

$$16 \text{ cm s}^{-1} = \omega \times 4 \text{ cm}$$

$$\omega = 4 \text{ s}^{-1}$$

$$\text{Period } T = \frac{2\pi}{\omega} = \frac{2\pi}{4 \text{ s}^{-1}} = 1.571 \text{ s}$$

Example 1.5 A spring stretches 0.15 m when a 0.3 kg mass is hung from it. The spring is then stretched an additional 0.1 m from this equilibrium point and released. Determine (i) spring constant k (ii) the amplitude of the oscillation A (iii) the maximum velocity v_0 .

Solution: (i) From Eq. (1.2)

$$\begin{aligned} k &= \frac{F}{x} = \frac{mg}{x} = \frac{(0.3 \text{ kg})(9.8 \text{ m s}^{-2})}{0.15 \text{ m}} \\ &= 19.6 \text{ N m}^{-1} \end{aligned}$$

(ii) The spring is stretched 0.1 m from equilibrium and released, the amplitude $A = 0.1 \text{ m}$.

(iii) The maximum velocity v_0 is obtained when the mass passes the equilibrium position.

From Eq. (1.6), we get

$$v_0 = \omega A.$$

From Eq. (1.5)

$$\omega = \sqrt{\frac{k}{m}}.$$

$$\begin{aligned} \text{Hence, } v_0 &= A \sqrt{\frac{k}{m}} = 0.1 \text{ m} \sqrt{\frac{19.6 \text{ N m}^{-1}}{0.3 \text{ kg}}} \\ &= 0.81 \text{ m s}^{-1} \end{aligned}$$

Example 1.6 The visible region of the electromagnetic spectrum is 400 nm to 700 nm. Calculate the frequency equivalent of the visible region in Hz. Velocity of light is $3 \times 10^8 \text{ m/s}$.

Solution: The frequency ν is given by

$$\nu = \frac{\text{Velocity}}{\text{Wavelength}}$$

When electromagnetic spectrum is 400 nm, we get

$$\nu = \frac{3 \times 10^8 \text{ m/s}}{400 \times 10^{-9} \text{ m}} = 7.5 \times 10^{14} \text{ Hz}$$

26 Engineering Physics

When electromagnetic spectrum is 700 nm, we get

$$v = \frac{3 \times 10^8 \text{ m/s}}{700 \times 10^{-9} \text{ m}} = 4.29 \times 10^{14} \text{ Hz}$$

The frequency equivalent of 400 nm to 700 nm is 7.5×10^{14} Hz to 4.29×10^{14} Hz.

Example 1.7 The displacement of a sound wave is given by

$$u(x, t) = 1.5 \times 10^{-3} \sin\left(\frac{2\pi x}{8} - 80\pi t\right)$$

where x is measured in metres and t in seconds. Evaluate the amplitude, wavelength, frequency and velocity of the wave.

Solution: The standard equation for the displacement is given by

$$u = A \sin 2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right)$$

The given equation can be written as

$$u = 1.5 \times 10^{-3} \sin 2\pi\left(\frac{x}{8} - 40t\right)$$

Comparing the two equations, we get

$$\text{Amplitude } (A) = 1.5 \times 10^{-3} \text{ m}, \quad \text{Wavelength } (\lambda) = 8 \text{ m},$$

$$\text{Period } T = \frac{1}{40} \text{ s}, \quad \text{Frequency } v = 1/T = 40 \text{ Hz}$$

$$\text{Velocity } v = v\lambda = (40 \text{ Hz}) (8 \text{ m}) = 320 \text{ m/s}$$

Example 1.8 Find the equation of a wave of amplitude 2 cm, period 0.5 s and velocity 200 cm/s moving along the x -axis.

Solution: Amplitude (A) = 2 cm, Frequency $= \frac{1}{0.5 \text{ s}} = 2 \text{ Hz}$

$$\text{Wavelength } (\lambda) = \frac{\text{Velocity}}{\text{Frequency}} = \frac{200 \text{ cm/s}}{2 \text{ Hz}} = 100 \text{ cm}$$

The general equation of a wave is given by

$$u = A \sin 2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right)$$

The equation of the wave having the given parameters is

$$u = 2 \sin 2\pi\left(\frac{x}{100} - \frac{t}{0.5}\right)$$

where x and y are in centimetres and t in seconds.

Example 1.9 A wave of wavelength 0.30 m is travelling down a 300 m long wire whose total mass is 15 kg. If the wire is under tension of 1000 N, what is the velocity and frequency of the wave?

Solution: Velocity of the wave in the string

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

where T is tension, m is the mass per unit length

$$v = \sqrt{\frac{\frac{1000 \text{ N}}{(15 \text{ kg})}}{(300 \text{ m})}} = \sqrt{\frac{(1000 \text{ N})(300 \text{ m})}{15 \text{ kg}}} = 141.4 \text{ m/s}$$

The frequency

$$v = \frac{v}{\lambda} = \frac{141.4 \text{ m/s}}{0.3 \text{ m}} = 471.3 \text{ Hz}$$

REVIEW QUESTIONS

- 1.1 When do you say that a motion is simple harmonic? Write the equation of motion of a particle of mass m executing simple harmonic motion.
- 1.2 Is the acceleration of a simple harmonic oscillator ever zero? If so, where?
- 1.3 If we double the amplitude of a SHM, how does this change the frequency, maximum velocity and total energy?
- 1.4 Write the differential equation of a damped harmonic oscillator whose damping force is proportional to its velocity. When do you say the system is critically damped?
- 1.5 Distinguish between critically damped and dead beat motions when a harmonic oscillator is damped by a force proportional to its velocity.
- 1.6 What do you understand by quality factor? On what factors does it depend?
- 1.7 Define wavelength and frequency of a periodic motion. State the relation between the two. Write the equation that represents periodic motion.
- 1.8 With examples, distinguish between transverse and longitudinal waves.
- 1.9 Explain (i) wave velocity and (ii) particle velocity. State the relation between the two.
- 1.10 For a wave travelling in the positive x -direction the displacement y is represented by a function of $(x - vt)$ as $y = f(x - vt)$. Show that v is the wave velocity of the wave.
- 1.11 Explain the difference between the speed of a transverse wave travelling down a rope and the speed of a tiny piece of the rope.
- 1.12 What kind of waves do you think will travel down a horizontal metal rod if you strike its end (i) vertically from above (ii) horizontally parallel to its length?

- 1.13** AM radio signals can usually be heard behind a hill, but FM often cannot. Explain.
(Hint: AM carrier wavelength is 200 to 600 m and for FM only about 3 m).
- 1.14** Explain how the amplitude and intensity of a wave changes with distance.
- 1.15** Show that the energy transported by a wave is proportional to the square of the amplitude.
- 1.16** What are the surface waves? Are they longitudinal? Are they transverse? Explain.
- 1.17** What do you understand by the phase of a wave? Two waves are represented by $y_1 = a \sin(\omega t - kx)$ and $y_2 = b \cos(\omega t - kx)$. What is the phase difference between the two?
- 1.18** What are spherical and plane waves? Give their analytical representations.
- 1.19** Write the equation of motion for a body of mass m executing SHM and solve the same. Show that the energy of a simple harmonic oscillator is proportional to the square of its amplitude.
- 1.20** Write the differential equation of a damped harmonic oscillator whose damping force is proportional to its velocity and obtain the solution for the same.
- 1.21** A body executing damped SHM is subjected to an oscillating external force. Assuming a damping force proportional to its velocity. Write its equation of motion and obtain its solution. Write a note on resonance.
- 1.22** Show that the expression $y = A \sin 2\pi \left(\frac{x}{\lambda} - \frac{t}{T} \right)$ represents a wavetrain of amplitude A , wavelength λ and period T , moving along the positive x -direction.
- 1.23** Show that the particle velocity at a point due to a passing wavetrain is equal to wave velocity times the slope of the displacement curve at that point. Derive the differential equation for wave motion in one dimension.
- 1.24** Derive an expression for the velocity of transverse vibrations in a stretched string.
- 1.25** State the three dimensional wave equation and solve the same to obtain the form of the displacement at (x, y, z, t) .

PROBLEMS

1. A body having a mass 4 g executes SHM. If the period is 0.2 s, find the force acting on the body when displacement is 4 cm.
2. The displacement of a particle executing SHM is given by $x = 10 \sin(0.2t + \phi)$. At $t = 0$, the particle is displaced by 1 cm. Determine (i) the initial phase (ii) the phase angle corresponding to 2 cm and (iii) the period.
3. A mass m at the end of a spring vibrates with a frequency of 0.88 Hz. When an additional 600 g mass is added to m , the frequency is 0.60 Hz. What is the value of m ?
4. Radio signals have frequencies between 550 kHz and 1600 kHz and travel with a speed of 3×10^8 m/s. What are the wavelengths of these signals?

5. A rope of mass 0.55 kg is stretched between two supports 30 m apart. If the tension in the rope is 150 N , how long will a pulse take to travel from one support to the other?
6. The particle displacement due to a sound wave is given by $y = 2 \times 10^{-3} \sin 2\pi \left(\frac{x}{5} - 50t \right)$ where x is measured in metres and t in seconds. What is the amplitude, wavelength, frequency, period and velocity of the wave?
7. The equation of a wave travelling in a string is given by $y = \cos \pi (0.5x - 200t)$ where x is in centimetres and t is in seconds. Evaluate the amplitude, wavelength, frequency, period and velocity of propagation.
8. Find the equation of a progressive wave of amplitude 0.02 m , frequency 550 Hz and velocity 330 m/s moving along the x -axis.
9. Two waves travelling along a stretched string have the same frequency, but one transports three times the power of the other. What is the ratio of the amplitudes of the two waves?
10. A wood piece on the surface of a pond is observed to move up and down a total vertical distance of 0.10 m , as a wave passes.
- What is the amplitude of the wave?
 - If the ripple increases to 0.15 m , what is the ratio of energy transported by the two waves?
11. A wave travelling along the x -axis is represented by

$$u(x, t) = A \sin 2\pi \left(\frac{x}{\lambda} - \frac{t}{T} \right)$$

Show that the magnitude of the velocity of a particle of the medium $|v_p| = v \frac{du}{dx}$ where v is the phase velocity of the wave.

C H A P T E R

2

Electromagnetic Theory

Sir Isaac Newton's corpuscular theory of light (1680) was the first scientific attempt to study the nature of light radiation. According to this concept, light consists of tiny elastic particles called **corpuscles** that travel in straight lines with the velocity of light. However, the theory could not account for phenomena such as interference, diffraction and polarization. It predicted a higher velocity in a denser medium than in a rarer medium. The theory that light is propagated as a wave through a hypothetical elastic medium called **ether**, developed by Huygens', Young and Fresnel, got established as it could explain most of the experimentally observed phenomena. At this stage, Clark Maxwell proposed (1864) his electromagnetic theory of light, which unified all the phenomena of electricity, magnetism and light into one magnificent theory. He established that the electric and magnetic fields are related to the charge and current densities by Maxwell's electromagnetic equations. While deriving the equations he found that Ampere's law is inconsistent in time varying fields. To make it consistent he introduced a quantity known as **displacement current**. The important result of Maxwell's equations was the prediction of the existence of electromagnetic waves and its velocity is proved to be same as that of light. Flow of energy or power in an electromagnetic field was also calculated using Poynting's theorem. To start with, we shall discuss some of the basic definitions in vector analysis.

2.1 THE DEL OPERATOR

The differential vector operator ∇ called **del or nabla** has many important applications in physical problems. It is capable of differentiating both vector and scalar functions with equal ease. It is defined as:

$$\nabla = \hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z} \quad (2.1)$$

where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are the unit vectors along the cartesian co-ordinate axes.

A differential operator can be treated in much the same way as any ordinary quantity. There are three possible operations with ∇ corresponding to the three possible types of vector multiplication and each one has its own physical significance. They are gradient, divergence and curl. We shall consider them one by one.

2.2 GRADIENT OF A SCALAR FUNCTION

If the vector differential operator ∇ is operated on a scalar function $\phi(x, y, z)$, this operation is known as gradient of a scalar function and abbreviated as Grad ϕ . It is defined as

$$\text{Grad } \phi = \nabla \phi = \hat{\mathbf{i}} \frac{\partial \phi}{\partial x} + \hat{\mathbf{j}} \frac{\partial \phi}{\partial y} + \hat{\mathbf{k}} \frac{\partial \phi}{\partial z} \quad (2.2)$$

The grad ϕ is a vector quantity. A familiar example is the electric field vector \mathbf{E} which is the negative gradient of electric potential V .

$$\mathbf{E} = -\text{Grad } V = -\nabla V \quad (2.3)$$

2.2.1 Physical Significance

Let scalar function ϕ represents temperature, then $\nabla \phi$ represents temperature gradient or rate of change of temperature with distance. Thus, although temperature is a scalar quantity having magnitude only, the temperature gradient $\nabla \phi$ is a vector quantity having both magnitude and direction. The direction of the resultant is the direction in which rate of change of temperature is maximum. This gradient vector can be expressed in the direction of co-ordinate axes x , y and z in terms of its components $\partial \phi / \partial x$, $\partial \phi / \partial y$, and $\partial \phi / \partial z$ and the resultant temperature gradient is the vector sum of these components given by Eq. (2.2). Thus, the gradient of a scalar function gives the particular direction in which the scalar function changes most rapidly or the maximum space rate of change of that function.

2.3 DIVERGENCE OF A VECTOR

The second type of del operation is the multiplication of del operator by another vector so that the resultant is the dot product or scalar product. The operation is called the **divergence of a vector \mathbf{A}** and is given by,

$$\begin{aligned} \text{Div } \mathbf{A} = \nabla \cdot \mathbf{A} &= \left(\hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z} \right) \cdot (\hat{\mathbf{i}} A_x + \hat{\mathbf{j}} A_y + \hat{\mathbf{k}} A_z) \\ &= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \end{aligned} \quad (2.4)$$

where A_x , A_y and A_z are the components of vector A in the direction of x , y and z , respectively.

2.3.1 Physical Significance

Vector analysis finds great usefulness in simplifying the expressions of the relations that exist in three dimensional fields. A consideration of fluid motion gives a direct interpretation of divergence. Consider the case of an incompressible fluid, say water. As in Fig. 2.1, the rectangular parallelopiped of dimension Δx , Δy , Δz is an infinitesimal volume element within the fluid.

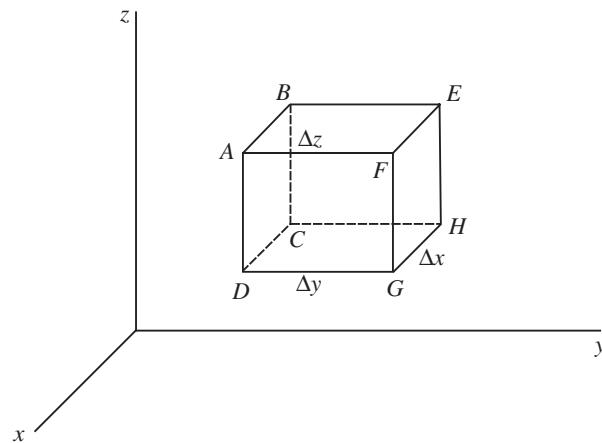


Fig. 2.1 A rectangular parallelopiped of dimension Δx , Δy , Δz .

If v_y is the average of the y -component of the fluid velocity through the left hand face, ABCD, then the flow into the parallelopiped through this face is $\rho_m v_y \Delta x \Delta z$, where ρ_m is the mass density of the fluid.

The rate of flow out through the face EFGH

$$= \left[\rho_m v_y + \frac{\partial(\rho_m v_y)}{\partial y} \Delta y \right] \cdot \Delta x \Delta z$$

The net outward flow in the y -direction in unit time = $\frac{\partial(\rho_m v_y)}{\partial y} \Delta x \Delta y \Delta z$

Similarly,

The net outward flow in the z -direction = $\frac{\partial(\rho_m v_z)}{\partial z} \Delta x \Delta y \Delta z$

The net outward flow in the x -direction = $\frac{\partial(\rho_m v_x)}{\partial x} \Delta x \Delta y \Delta z$

Considering all three directions, the net outflow in unit time is

$$\left[\frac{\partial(\rho_m v_x)}{\partial x} + \frac{\partial(\rho_m v_y)}{\partial y} + \frac{\partial(\rho_m v_z)}{\partial z} \right] \Delta x \Delta y \Delta z$$

The net outward flow per unit volume per unit time

$$= \frac{\partial(\rho_m v_x)}{\partial x} + \frac{\partial(\rho_m v_y)}{\partial y} + \frac{\partial(\rho_m v_z)}{\partial z} = \text{Div}(\rho_m \mathbf{v}) = \nabla \cdot (\rho_m \mathbf{v}) \quad (2.5)$$

This is the divergence of the fluid at any point x, y, z . At any point if outflow is greater than inflow, the divergence is said to be positive at that point. Whereas if outflow is less than inflow, divergence is negative. Hence for an incompressible fluid such as water, outflow is equal to inflow and the divergence is zero.

2.4 CURL OF A VECTOR

Another important operation with the vector operator ∇ is to cross it with a vector. The curl of a vector \mathbf{A} is defined by

$$\text{Curl } \mathbf{A} = \nabla \times \mathbf{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{\mathbf{k}} \quad (2.6)$$

This may be written conveniently in the determinant form,

$$\nabla \times \mathbf{A} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}$$

It is evident that the curl of a vector is also a vector.

2.4.1 Physical Significance

The concept of curl or rotation of a vector quantity is clearly illustrated in the stream flow problems. Figure 2.2 shows a stream on the surface of which floats a leaf in the $x-y$ plane.

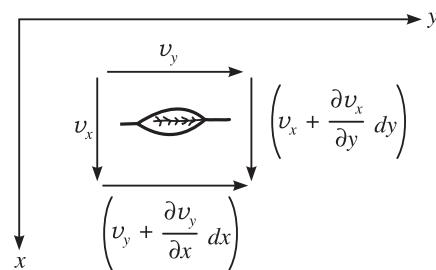


Fig. 2.2 A leaf floats in the $x-y$ plane.

34 Engineering Physics

There will be only a translational motion for the leaf, if the velocity at the surface is entirely in the y -direction and is uniform over the surface. But in general there will be rotational as well as translational motion because of the disturbances in the stream flow. Curl actually refers to the rotation of a vector quantity. The rate of rotation at any point is a measure of the curl of the velocity of water at that point. If the rotation is about the z -axis, the curl of velocity in the z -direction is denoted by $(\nabla \times \mathbf{v})_z$. A rotation from y to x is clockwise, whereas rotation from x to y is anticlockwise. From Fig. 2.2 it is evident that a positive value for $(\partial V_x / \partial y)$ will tend to rotate the leaf in clockwise direction, whereas a positive value of $(\partial V_y / \partial x)$ will tend to produce an anticlockwise rotation. Therefore, the rate of rotation about the z -axis is proportional to $(\partial v_y / \partial x - \partial v_x / \partial y)$ assuming that the velocity at the top is more. That is,

$$(\nabla \times \mathbf{v})_z = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}$$

Similarly, the corresponding components of the curl in x and y directions are

$$(\nabla \times \mathbf{v})_x = \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}$$

$$(\nabla \times \mathbf{v})_y = \frac{\partial v_z}{\partial x} - \frac{\partial v_x}{\partial z}$$

A rotation about any axis can always be expressed as the sum of the component rotations about x , y and z axes. Hence, the resultant rate of rotation or angular velocity is given by,

$$(\nabla \times \mathbf{v}) = \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \hat{\mathbf{k}} \quad (2.7)$$

The direction of the resultant curl is the axis of rotation. If the curl of a vector is zero, then the vector is said to be **irrotational**.

2.4.2 Important Vector Identities

Some of the important vector identities which are very useful in electromagnetic theory are listed as follows:

$$(i) \text{ Div} (\text{Grad } \phi) = \nabla \cdot (\nabla \phi) = \nabla^2 \phi \quad (2.8)$$

$$(ii) \text{ Curl} (\text{Grad } \phi) = \nabla \times \nabla \phi = 0 \quad (2.8a)$$

$$(iii) \text{ Div} (\text{Curl } \mathbf{A}) = \nabla \cdot (\nabla \times \mathbf{A}) = 0 \quad (2.8b)$$

$$(iv) \text{ Curl} (\text{Curl } \mathbf{A}) = \nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \quad (2.8c)$$

$$(v) \text{ Div} (\mathbf{A} \times \mathbf{B}) = \nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B}) \quad (2.8d)$$

2.4.3 Integral Theorems

There are two important integral theorems which will be of importance in electromagnetic theory — Gauss theorem and Stokes theorem.

Gauss theorem

If s is a closed surface surrounding the volume V , the divergence for any vector \mathbf{A} is expressed as,

$$\int_s \mathbf{A} \cdot d\mathbf{s} = \int_V (\nabla \cdot \mathbf{A}) dV \quad (2.9)$$

where $d\mathbf{s} = \hat{\mathbf{n}} d\mathbf{s}$, $d\mathbf{s}$ an element of area on s and $\hat{\mathbf{n}}$ the unit outward normal to s .

Stokes theorem

If C is the closed contour around the edge of the open surface S , then Stokes theorem states that

$$\int_l \mathbf{A} \cdot d\mathbf{l} = \int_s (\nabla \times \mathbf{A}) \cdot d\mathbf{s} \quad (2.10)$$

2.5 CONDUCTION AND DISPLACEMENT CURRENTS

We are familiar with conduction current. But in electromagnetic theory, we have the magnetic field produced due to time-varying electric field leading to the displacement current. In general, current density is defined as current per unit area.

2.5.1 Conduction Current

Conduction current is due to drift of electric charges in a conductor when an electric field is applied. Let \mathbf{E} be the electric field strength applied across a linear conductor of length l and cross-sectional area A . A current i flows through the conductor and V be the potential difference between the end points of the conductor. Then,

$$R = \frac{l}{\sigma A} \quad (2.11)$$

where R is the resistance between the points a and b and σ the conductivity. Applying Ohms law, we get

$$V = El = iR \quad (2.12)$$

Replacing R using Eq. (2.11), we get

$$\mathbf{E} = \frac{i}{\sigma A} = \frac{\mathbf{J}_1}{\sigma}$$

where

$$\mathbf{J}_1 = \frac{i}{A} = \sigma \mathbf{E} \quad (2.13)$$

\mathbf{J}_1 is the conduction current per unit area referred to as the **conduction current density** and is directly proportional to the electric field intensity.

2.5.2 Displacement Current

Consider the case of a capacitor of capacitance C , charged by applying a voltage V across its ends. Then, the current through the capacitor is

$$\mathbf{i}_c = \frac{dQ}{dt} = C \frac{dV}{dt} \quad (2.14)$$

where Q is the charge on the capacitor plates. In the case of a parallel plate capacitor,

$$C = \frac{\epsilon A}{d} \quad (2.15)$$

where A is the cross-sectional area of the plates, d the separation between the plates and ϵ the dielectric constant of the medium. Electric field strength

$$\mathbf{E} = \frac{V}{d} \quad (2.16)$$

Substituting Eq (2.15) and (2.16) in Eq. (2.14),

$$\begin{aligned} \mathbf{i}_c &= \epsilon A \frac{d\mathbf{E}}{dt} \\ \mathbf{J}_2 &= \frac{\mathbf{i}_c}{A} = \epsilon \frac{d\mathbf{E}}{dt} = \frac{d\mathbf{D}}{dt} \quad ; \quad \mathbf{D} = \epsilon \mathbf{E} \end{aligned} \quad (2.17)$$

where \mathbf{D} is the electric displacement density and \mathbf{J}_2 is called **displacement current density**, representing the current which directly passes through the capacitor and is only an apparent current representing the rate at which flow of charge takes place from electrode-to-electrode in the external circuit. Hence, it is known as **displacement current**. In an electromagnetic field, both conduction current and displacement current are present, hence the expression for total current density is,

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 = \sigma \mathbf{E} + \frac{d\mathbf{D}}{dt} \quad (2.18)$$

The concept of displacement current is very helpful in explaining the magnetic fields in free space due to time varying electric fields.

2.6 THE EQUATION OF CONTINUITY FOR TIME VARYING FIELDS

Since current is simply charge in motion, the total current flowing out of some volume must be equal to the rate of decrease of charge within the volume, assuming that charge cannot be created or destroyed. In mathematical terms, this conservation of charge concept can be stated as

$$\int_s \mathbf{J} \cdot d\mathbf{s} = -\frac{d}{dt} \int_V \rho dV \quad (2.19)$$

where ρ is the charge density defined as charge per unit volume. If the region of integration is stationary, Eq. (2.19) becomes,

$$\int_s \mathbf{J} \cdot d\mathbf{s} = - \int_V \frac{\partial \rho}{\partial t} dV$$

Converting surface integral to volume integral using Gauss theorem, we get

$$\begin{aligned} \int_V \nabla \cdot \mathbf{J} dV &= - \int_V \frac{\partial \rho}{\partial t} dV \\ \nabla \cdot \mathbf{J} &= - \frac{\partial \rho}{\partial t} \end{aligned} \quad (2.20)$$

This is the equation of continuity for time varying fields.

2.7 MAXWELL'S EQUATIONS

The electromagnetic wave phenomena are governed by a set of four equations collectively known as **Maxwell's field equations**, since Maxwell formulated them in a general way as a self-consistent set.

2.7.1 Maxwell's Electric Flux Equation

According to Gauss's law in electrostatics, the total electric displacement through the surface enclosing a volume is equal to the total charge within the volume:

$$\int_s \mathbf{D} \cdot d\mathbf{s} = q \quad (2.21)$$

where \mathbf{D} is the electric displacement density and q , the electric charge. In terms of charge density ρ ,

$$q = \int_V \rho dV \quad (2.22)$$

From Eqs. (2.21) and (2.22)

$$\int_s \mathbf{D} \cdot d\mathbf{s} = \int_V \rho dV$$

Converting the surface integral to volume integral,

$$\begin{aligned} \int_V \nabla \cdot \mathbf{D} dV &= \int_V \rho dV \\ \nabla \cdot \mathbf{D} &= \rho \end{aligned} \quad (2.23)$$

Equation (2.23) is referred to as **Maxwell's electric flux equation**.

2.7.2 Maxwell's Magnetic Flux Equation

According to Gauss theorem in magnetostatics, the net magnetic flux emerging through any closed surface is zero.

$$\int_s \mathbf{B} \cdot d\mathbf{s} = 0 \quad (2.24)$$

where \mathbf{B} is the magnetic flux density which is magnetic flux per unit area. Converting the surface integral of Eq. (2.24) to volume integral, we get

$$\begin{aligned} \int_V \nabla \cdot \mathbf{B} dV &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned} \quad (2.25)$$

Equation (2.25) is referred to as **Maxwell's magnetic flux equation**.

2.7.3 Maxwell's Equation for Induced Electric Field

According to Faraday's law of electromagnetic induction, the negative time rate of change of magnetic flux linked with a circuit is equal to the emf induced in the circuit.

$$\oint_l \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial}{\partial t} \int_s \mathbf{B} \cdot d\mathbf{s} \quad (2.26)$$

where \mathbf{E} is the electric field strength and \mathbf{B} is the magnetic flux density. Converting the line integral to surface integral using Stokes theorem,

$$\int_s (\nabla \times \mathbf{E}) \cdot d\mathbf{s} = \int_s -\frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{s} \quad (2.27)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.28)$$

Equation (2.28) is referred to as **Maxwell's equation for induced electric field**.

2.7.4 Maxwell's Equation for Induced Magnetic Field

According to Ampere's circuital law, the magnetomotive force around a closed path is equal to the current enclosed by the path. Mathematically, we obtain

$$\int_l \mathbf{H} \cdot d\mathbf{l} = i \quad (2.29)$$

where \mathbf{H} is the magnetic field strength and i the current enclosed by the path. If \mathbf{J} is the current density,

$$i = \int_s \mathbf{J} \cdot d\mathbf{s} \quad (2.30)$$

From Eqs. (2.29) and (2.30), we get

$$\int_l \mathbf{H} \cdot d\mathbf{l} = \int_s \mathbf{J} \cdot ds \quad (2.31)$$

Again converting the line integral to surface integral

$$\begin{aligned} \int_s \nabla \times \mathbf{H} \cdot ds &= \int_s \mathbf{J} \cdot ds \\ \nabla \times \mathbf{H} &= \mathbf{J} \end{aligned} \quad (2.32)$$

Equation (2.32) is the differential form of Ampere's circuital law. Taking the divergence of Ampere's circuital law and using Eq. (2.8c), yields the equation of continuity for steady currents:

$$\begin{aligned} \nabla \cdot (\nabla \times \mathbf{H}) &= \nabla \cdot \mathbf{J} = 0 \\ \nabla \cdot \mathbf{J} &= 0 \end{aligned} \quad (2.33)$$

Thus, Ampere's law is not consistent with the time-varying equation of continuity, Eq. (2.20). Hence, Maxwell in the mid-1860's modified Ampere's law. Substituting, Maxwell's electric flux equation, Eq. (2.23) into the equation of continuity, (Eq. 2.20), we get

$$\nabla \cdot \mathbf{J} = - \frac{\partial(\nabla \cdot \mathbf{D})}{\partial t}$$

Interchanging the differentiations with respect to space and time and rearranging

$$\nabla \cdot \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \right) = 0 \quad (2.34)$$

This may be put in the integral form by integrating over a volume and then applying the divergence theorem, we get

$$\int_s \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \right) \cdot ds = 0 \quad (2.35)$$

Here $(\partial \mathbf{D} / \partial t + \mathbf{J})$ may be regarded as the total current density for time varying fields. Since \mathbf{D} is the displacement density, $\partial \mathbf{D} / \partial t$ is known as the **displacement current density**. Maxwell reasoned that the total current density should replace \mathbf{J} in Ampere's law with the result that

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \quad (2.36)$$

Integration of Eq. (2.36) over a surface and application of Stokes's theorem leads to the integral form

$$\oint_l \mathbf{H} \cdot d\mathbf{l} = \int_s \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \right) \cdot ds \quad (2.37)$$

Equation (2.36) is referred as **Maxwell's equation for induced magnetic field**. Equations (2.23), (2.25), (2.28) and (2.36) together form the Maxwell's equations.

2.8 ELECTROMAGNETIC WAVES IN DIELECTRIC MEDIUM

The fundamental relations that must be satisfied are the four Maxwell's equations, (2.23), (2.25), (2.28) and (2.36). In addition there are three constitutive relations that concern the characteristics of the medium in which the fields exist. They are

$$\mathbf{D} = \epsilon \mathbf{E} ; \quad \mathbf{B} = \mu \mathbf{H} ; \quad \mathbf{J} = \sigma \mathbf{E}$$

where ϵ , μ and σ are the permittivity, permeability and conductivity of the medium which is assumed to be homogeneous, isotropic and source free. A homogeneous medium is one for which ϵ , μ and σ are constants throughout the medium. The medium is isotropic if ϵ is a scalar constant, so that \mathbf{D} and \mathbf{E} possess the same direction. Source free regions are regions in which there are no impressed voltages or currents. In free space, the constitutive relations become,

$$\mathbf{J} = \sigma \mathbf{E} = 0 ; \quad \mathbf{D}_0 = \epsilon_0 \mathbf{E} ; \quad \mathbf{B}_0 = \mu_0 \mathbf{H}$$

where μ_0 and ϵ_0 are the permeability and permittivity of free space. Using these results, Maxwell's equations can be simplified as follows:

$$\nabla \cdot \mathbf{E} = 0 \tag{2.38}$$

$$\nabla \cdot \mathbf{H} = 0 \tag{2.39}$$

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{2.40}$$

$$\nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{2.41}$$

Equations (2.38) to (2.41) are **Maxwell's equations in free space**. Taking the curl of Eq. (2.40), we get

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left[-\mu_0 \frac{\partial \mathbf{H}}{\partial t} \right] \tag{2.42}$$

Using the identity in Eq. (2.8d), Eq. (2.42) becomes,

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \nabla \times \frac{\partial \mathbf{H}}{\partial t}$$

But from Eq. (2.38) $\nabla \cdot \mathbf{E} = 0$. Also interchanging the space and time derivatives, we get

$$\nabla^2 \mathbf{E} = \mu_0 \frac{\partial(\nabla \times \mathbf{H})}{\partial t} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{2.43}$$

where we have used Eq. (2.41). Proceeding on similar lines, we obtain

$$\nabla^2 \mathbf{H} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{H}}{\partial t^2} \tag{2.44}$$

These equations resemble the general wave equation, Eq. (1.82), which reveals that **E** and **H** propagate as waves through free space. Equations (2.43) and (2.44) are the wave equations for **E** and **H** and they propagate with a speed,

$$v = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \quad (2.45)$$

Substituting the values of μ_0 and ϵ_0 , Maxwell obtained the velocity as 3.11×10^8 m/s which was in agreement with the velocity of light measured. Hence, Maxwell predicted that light is electromagnetic in nature and travels with a velocity of 3×10^8 m/s in free space.

Instead, if we consider electromagnetic waves in a medium, its velocity would be

$$v = \frac{1}{\sqrt{\mu \epsilon}} \quad (2.45a)$$

where μ is the permeability of the medium and ϵ , the permittivity of the medium. The relative permeability μ_r and relative permittivity ϵ_r are defined by

$$\mu = \mu_0 \mu_r \quad \text{and} \quad \epsilon = \epsilon_0 \epsilon_r \quad (2.45b)$$

2.9 TRANSVERSE NATURE OF ELECTROMAGNETIC WAVES

The wave equations satisfied by the electric and magnetic fields in free space are given by Eqs. (2.43) and (2.44). A uniform plane wave is the one in which **E** and **H** lie in a plane and have the same value everywhere in that plane at any fixed instant. For convenience, consider a uniform plane wave propagating in x – direction in which **E** and **H** depends only on x . Then the derivatives with respect to y and z vanish. The wave equations then reduce to,

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (2.46)$$

$$\frac{\partial^2 \mathbf{H}}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (2.47)$$

In terms of the components, Eq. (2.46) reduces to the three equations as follows:

$$\frac{\partial^2 E_x}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 E_x}{\partial t^2} \quad (2.48)$$

$$\frac{\partial^2 E_y}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 E_y}{\partial t^2} \quad (2.49)$$

$$\frac{\partial^2 E_z}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 E_z}{\partial t^2} \quad (2.50)$$

42 Engineering Physics

As we know from Eq. (2.38) that

$$\frac{\partial E_x}{\partial x} = 0 \quad (2.51)$$

\therefore Equation (2.48) reduces to

$$\frac{\partial^2 E_x}{\partial t^2} = 0 \quad (2.52)$$

This implies that E_x be either zero, constant in time, or increasing uniformly with time. A field satisfying either of the last two of these conditions would not be a part of the wave motion, and so the first condition must be satisfied. Therefore, a uniform plane wave progressing in the x -direction has no x component of \mathbf{E} . A similar analysis would show that there is no x -component of \mathbf{H} also. It follows, therefore, that uniform plane electromagnetic waves are transverse and have components of \mathbf{E} and \mathbf{H} only in directions perpendicular to the direction of propagation.

2.10 RELATION BETWEEN \mathbf{E} AND \mathbf{H}

For a uniform plane wave propagating in the x -direction \mathbf{E} and \mathbf{H} are both independent of y and z and also $E_x = H_x = 0$. Consequently,

$$\nabla \times \mathbf{E} = - \frac{\partial E_z}{\partial x} \hat{\mathbf{y}} + \frac{\partial E_y}{\partial x} \hat{\mathbf{z}}$$

$$\nabla \times \mathbf{H} = - \frac{\partial H_z}{\partial x} \hat{\mathbf{y}} + \frac{\partial H_y}{\partial x} \hat{\mathbf{z}}$$

Then from the Maxwell's equations, Eq. (2.40) and Eq. (2.41)

$$-\frac{\partial E_z}{\partial x} \hat{\mathbf{y}} + \frac{\partial E_y}{\partial x} \hat{\mathbf{z}} = -\mu_0 \left(\frac{\partial H_y}{\partial t} \hat{\mathbf{y}} + \frac{\partial H_z}{\partial t} \hat{\mathbf{z}} \right) \quad (2.53)$$

$$-\frac{\partial H_z}{\partial x} \hat{\mathbf{y}} + \frac{\partial H_y}{\partial x} \hat{\mathbf{z}} = \epsilon_0 \left(\frac{\partial E_y}{\partial t} \hat{\mathbf{y}} + \frac{\partial E_z}{\partial t} \hat{\mathbf{z}} \right) \quad (2.54)$$

Equating the coefficients of $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ in Eqs. (2.53) and (2.54), yields four equations as follows:

$$\frac{\partial E_z}{\partial x} = \mu_0 \frac{\partial H_y}{\partial t} \quad (2.55)$$

$$\frac{\partial E_y}{\partial x} = -\mu_0 \frac{\partial H_z}{\partial t} \quad (2.56)$$

$$-\frac{\partial H_z}{\partial x} = \epsilon_0 \frac{\partial E_y}{\partial t} \quad (2.57)$$

$$\frac{\partial H_y}{\partial x} = \epsilon_0 \frac{\partial E_z}{\partial t} \quad (2.58)$$

The electric field vector for a plane wave that propagates along the positive direction of x -axis is a function of $(x - v_0 t)$. Hence,

$$E_y = f_1(x - v_0 t) ; v_0 = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

$$\frac{\partial E_y}{\partial t} = \frac{\partial f_1(x - v_0 t)}{\partial (x - v_0 t)} \frac{\partial (x - v_0 t)}{\partial t} = -v_0 \frac{\partial f_1(x - v_0 t)}{\partial (x - v_0 t)}$$

Substituting for $\partial E_y / \partial t$ in Eq. (2.57), we get

$$\frac{\partial H_z}{\partial x} = v_0 \epsilon_0 \frac{\partial f_1(x - v_0 t)}{\partial (x - v_0 t)} \quad (2.59)$$

$$\text{But } \frac{\partial f_1(x - v_0 t)}{\partial x} = \frac{\partial f_1(x - v_0 t)}{\partial (x - v_0 t)} \frac{\partial (x - v_0 t)}{\partial x} = \frac{\partial f_1(x - v_0 t)}{\partial (x - v_0 t)}$$

Hence, Eq. (2.59) takes the form

$$\frac{\partial H_z}{\partial x} = v_0 \epsilon_0 \frac{\partial f_1(x - v_0 t)}{\partial x} \quad (2.60)$$

Integrating Eq. (2.60) and substituting for v_0 , we get

$$\begin{aligned} H_z &= \sqrt{\frac{\epsilon_0}{\mu_0}} f_1(x - v_0 t) + C \\ H_z &= \sqrt{\frac{\epsilon_0}{\mu_0}} E_y + C \end{aligned} \quad (2.61)$$

Since H_z and E_y in Eq. (2.61) represents field parameters, C should also represent a field parameter, but it must be independent of x . For a wave propagating in x -direction, a field parameter independent of x cannot be a part of the wave motion. Hence, we can neglect the value of C . Therefore,

$$\frac{E_y}{H_z} = \sqrt{\frac{\mu_0}{\epsilon_0}} \quad (2.62)$$

Similarly, it can be shown that

$$\frac{E_z}{H_y} = -\sqrt{\frac{\mu_0}{\epsilon_0}} \quad (2.63)$$

44 Engineering Physics

Since,

$$E = |E| = \sqrt{E_y^2 + E_z^2} \quad \text{and} \quad H = |H| = \sqrt{H_y^2 + H_z^2} \quad (2.64)$$

$$\frac{E}{H} = \sqrt{\frac{\mu_0}{\epsilon_0}} \quad (2.64a)$$

where E and H are the total electric and magnetic field intensities and this definite ratio is denoted by η . This ratio is equal to the square root of permeability to dielectric constant of the non-conducting medium and is called **characteristic impedance** or **intrinsic impedance** of free space. It is called so, because E/H has the dimension of impedance or ohms. Substituting the values of μ_0 and ϵ_0 , we get

$$\eta = \sqrt{\frac{\mu_0}{\epsilon_0}} = 120\pi \approx 377 \text{ ohm} \quad (2.65)$$

Generalising Eq. (2.65) for a lossless medium we have

$$\eta = \frac{\mathbf{E}}{\mathbf{H}} = \sqrt{\frac{\mu}{\epsilon}} \quad (2.65a)$$

where μ and ϵ are the permeability and permittivity of the medium, respectively. Since the unit of \mathbf{E} is V/m and that of \mathbf{H} is A/m , the unit of η is that of impedance or ohm.

Using Eqs. (2.62) and (2.63), we get

$$\mathbf{E} \cdot \mathbf{H} = E_y H_z - E_z H_y = \eta H_y H_z - \eta H_y H_z = 0 \quad (2.66)$$

$$\mathbf{E} \times \mathbf{H} = \hat{x}(E_y H_z - E_z H_y) = \hat{x}\eta H^2 \quad (2.66a)$$

Thus, in a uniform plane wave, \mathbf{E} and \mathbf{H} are at right angles to each other, and the electric field vector crossed into the magnetic field vector gives the direction of propagation of the wave.

2.11 POYNTING'S THEOREM

Electromagnetic wave carries energy from source to receivers as they propagate through space. Poynting's theorem gives a relationship between rate of this energy transfer and the amplitude of electric and magnetic field strengths. It is actually the mathematical representation of law of conservation of energy applied to the electromagnetic fields. From Maxwell's equations Eq. (2.36), we get

$$\mathbf{J} = \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} \quad (2.67)$$

Taking dot product with \mathbf{E} on both sides,

$$\mathbf{E} \cdot \mathbf{J} = \mathbf{E} \cdot (\nabla \times \mathbf{H}) - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \quad (2.68)$$

Using the identity given in Eq. (2.8d), we get

$$\mathbf{E} \cdot \mathbf{J} = \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \nabla \cdot (\mathbf{E} \times \mathbf{H}) - \epsilon \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \quad (2.69)$$

Using the Maxwell's equation given by Eq. (2.28), we get

$$\mathbf{E} \cdot \mathbf{J} = -\mu \mathbf{H} \cdot \frac{\partial \mathbf{H}}{\partial t} - \epsilon \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} - \nabla \cdot (\mathbf{E} \times \mathbf{H}) \quad (2.70)$$

Since

$$\mathbf{H} \cdot \frac{\partial \mathbf{H}}{\partial t} = \frac{1}{2} \frac{\partial \mathbf{H}^2}{\partial t} \quad \text{and} \quad \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{2} \frac{\partial \mathbf{E}^2}{\partial t} \quad (2.71)$$

Equation (2.71) reduces to

$$\mathbf{E} \cdot \mathbf{J} = -\frac{\mu}{2} \frac{\partial H^2}{\partial t} - \frac{\epsilon}{2} \frac{\partial E^2}{\partial t} - \nabla \cdot (\mathbf{E} \times \mathbf{H})$$

Integrating over a volume V ,

$$\int_V \mathbf{E} \cdot \mathbf{J} dV = -\frac{\partial}{\partial t} \int_V \left(\frac{\mu H^2}{2} + \frac{\epsilon E^2}{2} \right) dV - \int_V \nabla \cdot (\mathbf{E} \times \mathbf{H}) dV \quad (2.72)$$

The last term in Eq. (2.72) is converted to surface integral using Gauss theorem,

$$\int_V \mathbf{E} \cdot \mathbf{J} dV = -\frac{\partial}{\partial t} \int_V \left(\frac{\mu H^2}{2} + \frac{\epsilon E^2}{2} \right) dV - \int_s (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{s} \quad (2.73)$$

A physical interpretation of each term in Eq. (2.73) gives an idea about Poynting's theorem.

$$\int_V \mathbf{E} \cdot \mathbf{J} dV$$

This term represents the power dissipated in the volume V . Power dissipated per unit volume of a conductor is $(V_i I / Ad) = EJ$ watts, where $\mathbf{E} = (V_i/d)$ is the voltage drop per unit length, I is the current through the conductor and A its cross sectional area. Hence, the total power dissipated in the volume V is $\int_V \mathbf{E} \cdot \mathbf{J} dV$.

$$-\frac{\partial}{\partial t} \int_V \left(\frac{\mu H^2}{2} + \frac{\epsilon E^2}{2} \right) dV$$

$\epsilon E^2/2$ and $\mu H^2/2$ together represent the energy density or the energy stored per unit volume in the dielectric medium in the form of electric and magnetic fields, respectively. The integral represents the stored energy in the volume V . The negative time derivative of this quantity then represents the rate at which the stored energy in the volume is decreasing.

$$-\int_s (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{s}$$

46 Engineering Physics

According to the law of conservation of energy, the rate of energy dissipation in the volume V must be equal to the rate at which the stored energy in the volume V is decreasing plus the rate at which energy is entering the volume V from outside. The term $-\int_s (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{s}$, therefore,

must represent the rate of flow of energy entering through the surface of the volume. Then the term $\int_s (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{s}$ represents the rate of flow of energy outward through the surface enclosing the volume.

The interpretation of Eq. (2.73) leads to the conclusion that the integral of $\mathbf{E} \times \mathbf{H}$ over any closed surface gives the rate of energy flow or the power per unit area through that surface. The vector,

$$\mathbf{P} = \mathbf{E} \times \mathbf{H} \quad (2.74)$$

represents the **Poynting vector**. The direction of flow is perpendicular to the plane containing \mathbf{E} and \mathbf{H} .

SOLVED EXAMPLES

Example 2.1 Determine the peak value of displacement current through a 10 pF capacitor if $V = 0.2 \sin 120 \pi t$ volt.

Solution: $V = 0.2 \sin 120 \pi t$

$$C = 10 \text{ pF} = 10 \times 10^{-12} \text{ F}$$

$$\begin{aligned} I &= C \frac{dV}{dt} = 10 \times 10^{-12} \frac{d}{dt}(0.2 \sin 120 \pi t) \\ &= 10 \times 10^{-12} \times 0.2 \times (120\pi) (\cos 120 \pi t) \end{aligned}$$

For the cosine function, peak value occurs when $120 \pi t = 0$. Hence,

$$I = 0.7536 \times 10^{-9} \text{ A}$$

Example 2.2 Calculate the peak value of displacement current density in a good conductor for which $\epsilon_r = 1$ when $\mathbf{E} = \sin 120 \pi t$ volt.

Solution: Displacement current density, $\mathbf{J}_2 = \epsilon \frac{d\mathbf{E}}{dt}$

$$\begin{aligned} \mathbf{J}_2 &= \epsilon_0 \epsilon_r \frac{d}{dt} (\sin 120 \pi t \text{ volt}) \\ &= (8.854 \times 10^{-12} \text{ F/m}) \times (120\pi) (\cos 120 \pi t) \\ &= 3.34 \times 10^{-9} \text{ A/m}^2 \end{aligned}$$

Example 2.3 Find displacement current density for the magnetic field given by, $H_x = 0$, $H_y = 0$ and $H_z = H_0 \sin(\omega t - \beta x)$. Assume the conduction current density to be zero.

Solution: From Eq. (2.32), we get

$$\text{Curl } \mathbf{H} = \mathbf{J}_1 + \mathbf{J}_2 = \mathbf{J}_2$$

$$\text{Curl } \mathbf{H} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & H_z \end{vmatrix}$$

$$\text{Curl } \mathbf{H} = \hat{\mathbf{i}} \left(\frac{\partial H_z}{\partial y} - 0 \right) - \hat{\mathbf{j}} \left(\frac{\partial H_z}{\partial x} - 0 \right) + \hat{\mathbf{k}} (0 - 0)$$

We have

$$\frac{\partial H_z}{\partial y} = 0$$

$$\frac{\partial H_z}{\partial x} = -H_0 \beta \cos(\omega t - \beta x)$$

Hence,

$$\mathbf{J}_2 = \hat{\mathbf{j}} H_0 \beta \cos(\omega t - \beta x)$$

Example 2.4 Calculate the value of Poynting vector for a 60 W bulb at a distance of 0.5 m from it.

$$\begin{aligned} \text{Solution: Poynting's vector, } \mathbf{P} &= \frac{\text{Power}}{\text{Area}} = \frac{60 \text{ W}}{4\pi (0.5 \text{ m})^2} \\ &= 19.1 \text{ W/m}^2 \end{aligned}$$

Example 2.5 A plane electromagnetic wave travelling in $+x$ direction in an unbounded lossless dielectric medium of $\mu_r = 1$ and $\epsilon_r = 3$ has a peak electric field intensity \mathbf{E} of 6 V/m. Find (i) the velocity of the wave (ii) the intrinsic impedance of the medium and (iii) the peak value of the magnetic intensity.

Solution: From Eq. (2.64)

$$\mathbf{E}_{\text{peak}} = \sqrt{E_y^2 + E_z^2} = 6 \text{ V/m}$$

From Eqs. (2.45a) and (2.45b), velocity is given by

$$v = \frac{1}{\sqrt{\mu\epsilon}} = \frac{1}{\sqrt{\mu_0\epsilon_0}} \cdot \frac{1}{\sqrt{\mu_r\epsilon_r}} = \frac{c}{\sqrt{\mu_r\epsilon_r}}$$

48 Engineering Physics

where c is the velocity of light

$$(i) v = \frac{3 \times 10^8 \text{ m/s}}{\sqrt{1 \times 3}} = 1.732 \times 10^8 \text{ ms}^{-1}$$

(ii) Intrinsic impedance

$$\begin{aligned}\eta &= \sqrt{\frac{\mu}{\epsilon}} = \sqrt{\frac{\mu_0}{\epsilon_0}} \sqrt{\frac{\mu_r}{\epsilon_r}} = 377 \times \frac{1}{\sqrt{3}} \\ &= 217.66 \Omega\end{aligned}$$

(iii) Peak value of the magnetic intensity

$$H_p = E_{\text{peak}} \sqrt{\frac{\epsilon}{\mu}} = \frac{6 \text{ V}}{217.66 \Omega} = 2.76 \times 10^{-2} \text{ Am}^{-1}$$

REVIEW QUESTIONS

- 2.1 Define gradient of a scalar function and divergence and curl of a vector.
- 2.2 Distinguish between conduction current density and displacement current density.
- 2.3 Define divergence of a vector. Explain its physical significance.
- 2.4 Define the curl of a vector. Explain its physical significance.
- 2.5 State and explain the equation of continuity for time varying fields.
- 2.6 What is displacement current density?
- 2.7 State the four Maxwell's equations.
- 2.8 Write the word statement of Maxwell's equations.
- 2.9 Deduce the Maxwell's equations for free space from the Maxwell's equations.
- 2.10 State and explain Poynting's theorem.
- 2.11 Show that Ampere's law is inconsistent in time varying fields.
- 2.12 Explain the physical significance of gradient, divergence and curl.
- 2.13 Derive Maxwell's equations.
- 2.14 Derive the free space electromagnetic wave equations and show that the velocity of the electromagnetic wave is $1/\sqrt{\mu_0 \epsilon_0}$.
- 2.15 Show that electromagnetic waves are transverse in nature and have components of \mathbf{E} and \mathbf{H} in directions perpendicular to the direction of propagation.

- 2.16** For a plane wave propagating along x -axis, obtain the relation connecting \mathbf{E} and \mathbf{H} . What is characteristic impedance?
- 2.17** State Maxwell's equations and obtain Poynting theorem.

PROBLEMS

1. A uniform plane wave propagating in free space is given by $\mathbf{E} = 500 \cos(\mathbf{kz} - 2 \times 10^{14} t)$ V/m. If the intrinsic impedance of free space is 120π ohms. Find the wave vector, velocity of propagation and the expression for \mathbf{H} .
2. Show that the ratio of amplitudes of conduction and displacement current density is $\frac{\sigma}{\omega\epsilon}$ for a capacitor, if the applied field is $\mathbf{E} = E_m \cos \omega t$.
3. The components of the electric field in free space are $E_z = E_0 \cos kx \sin \omega t$. Evaluate the magnetic field intensity \mathbf{H} .
4. Prove that the velocity of the electromagnetic wave in a dielectric medium is always less than the velocity in free space.
5. For a plane electromagnetic wave travelling along x -axis, show that $\frac{\partial^2 E_x}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 E_x}{\partial t^2}$. Using this differential equation, show that a sine wave in E_x travelling in the negative x -direction will have a velocity $= 1/\sqrt{\mu_0 \epsilon_0}$
6. If $\mathbf{E} = \hat{\mathbf{j}} E_0 \sin(kx + \omega t)$ and $\mathbf{H} = -\hat{\mathbf{k}} H_0 \sin(kx + \omega t)$, show that these fields can satisfy Maxwell's equations if ω and k are related in a certain way. Find the relation.

C H A P T E R

3

Interference

Isaac Newton (1680) was the first to propose a theory for the nature of light radiation. His corpuscular theory of light could not explain many of the common phenomena such as interference, diffraction and polarization. In 1678, the Dutch scientist Huygens proposed that light is propagated as a wave through a hypothetical all pervasive medium called **ether**. Huygen's principle gives a method for predicting the future position of a wave front when an earlier position is known. The wave theory got established through the work of Young and Fresnel. In this chapter, we discuss some of the interesting results of their interference experiments. The light distribution resulting from the superposition of waves is called **interference**.

3.1 SUPERPOSITION OF WAVES

According to Young (1801), when a particle in a medium is disturbed simultaneously by two or more waves, the resultant displacement in the medium at any instant will be the sum of the displacements due to each wave, separately. This is the superposition principle. If y_1 and y_2 are the displacements of the two waves, then the resultant displacement y is given by

$$y = y_1 + y_2 \quad (3.1)$$

When two or more waves superimpose, there will be positions of maximum and minimum intensities. This modification in the intensity of light, due to the superposition of two or more wave trains is called **interference**. If the intensity is increased, we call it **constructive interference** and if there is reduction in intensity we call it **destructive interference**.

3.1.1 Analytical Treatment

The prerequisite of interference is the superposition of waves. Consider two waves having the same frequency but differing in phase by a constant value of δ

$$y_1 = a_1 \sin \omega t \quad \text{and} \quad y_2 = a_2 \sin(\omega t + \delta) \quad (3.2)$$

Applying superposition principle given by Eq. (3.1), we get

$$\begin{aligned} y &= a_1 \sin \omega t + a_2 \sin(\omega t + \delta) \\ &= a_1 \sin \omega t + a_2 \sin \omega t \cos \delta + a_2 \cos \omega t \sin \delta \\ &= (a_1 + a_2 \cos \delta) \sin \omega t + (a_2 \sin \delta) \cos \omega t \end{aligned} \quad (3.3)$$

The terms in the parenthesis are constants in time. Introducing new constants A and ϕ by the relations

$$a_1 + a_2 \cos \delta = A \cos \phi \quad (3.4)$$

$$a_2 \sin \delta = A \sin \phi \quad (3.5)$$

Equation (3.3) is modified as

$$\begin{aligned} y &= A \cos \phi \sin \omega t + A \sin \phi \cos \omega t \\ &= A \sin(\omega t + \phi) \end{aligned} \quad (3.6)$$

Squaring and adding Eqs. (3.4) and (3.5), we get

$$A^2 = a_1^2 + a_2^2 + 2a_1 a_2 \cos \delta \quad (3.7)$$

Dividing Eq. (3.5) by Eq. (3.4), we obtain

$$\tan \phi = \frac{a_2 \sin \delta}{a_1 + a_2 \cos \delta} \quad (3.8)$$

Since intensity is proportional to the square of amplitude, taking the constant of proportionality as unity.

$$I = a_1^2 + a_2^2 + 2a_1 a_2 \cos \delta \quad (3.9)$$

Condition for maximum intensity

The intensity I , in Eq. (3.9), is maximum at points where $\cos \delta = 1$. That is when

$$\delta = 2n\pi, \quad n = 0, \pm 1, \pm 2, \dots \quad (3.10)$$

Since a path difference of λ is equivalent to a phase difference of 2π , the condition for maximum intensity reduces to

$$\text{Path difference} = n\lambda, \quad n = 0, 1, 2, \dots \quad (3.11)$$

The value of maximum intensity is

$$I_{\max} = a_1^2 + a_2^2 + 2a_1 a_2 = (a_1 + a_2)^2 \quad (3.12)$$

When $a_1 = a_2 = a$,

$$I_{\max} = 4a^2$$

Obviously, $I_{\max} > a_1^2 + a_2^2$. Thus, the resultant intensity at certain points is greater than the sum of the individual intensities, leading to constructive interference.

Condition for minimum intensity

I is minimum at points where $\cos \delta = -1$. That is, when

$$\delta = (2n - 1)\pi, \quad n = \pm 1, \pm 2, \dots$$

In terms of path difference, the condition for minimum intensity becomes

$$\text{Path difference} = (2n - 1) \frac{\lambda}{2}, \quad n = 1, 2, \dots \quad (3.13)$$

The value of minimum intensity is

$$I_{\min} = a_1^2 + a_2^2 - 2a_1a_2 = (a_1 - a_2)^2 \quad (3.14)$$

When $a_1 = a_2$,

$$I_{\min} = 0$$

Intensity distribution

At points that lie between maxima and minima, assuming that $a_1 = a_2 = a$, from Eq. (3.9) we have

$$\begin{aligned} I &= 2a^2(1 + \cos \delta) = 4a^2 \cos^2\left(\frac{\delta}{2}\right) \\ I &= 4I_0 \cos^2\left(\frac{\delta}{2}\right) \end{aligned} \quad (3.15)$$

where $I_0 = a^2$ is the intensity due to a single wave. The intensity distribution is shown in Fig. 3.1.

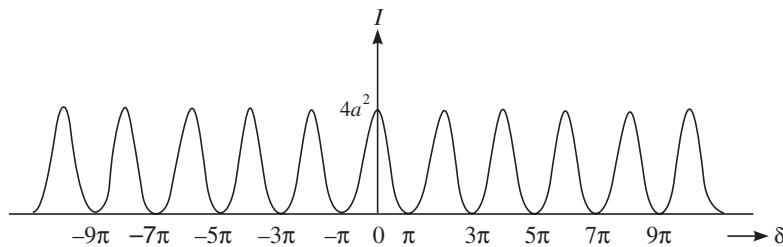


Fig. 3.1 A plot of the intensity of the interference pattern as a function of the phase difference between the waves.

The intensity distribution resulting from superposition of waves consists of alternate bright and dark bands called **interference fringes**. The phenomenon of interference, thus, causes energy to be redistributed.

During superposition, resultant intensity is increased at points where crest fall on crest and trough fall on trough, whereas the intensity is reduced where crest fall on trough and trough fall on crest. Thus, the phenomenon of interference, confirms Max Born's statement that light + light does not always give more light, but may in certain circumstances give darkness.

3.2 YOUNG'S DOUBLE SLIT EXPERIMENT

The important requirement for the production of a well defined interference pattern is that the light waves that come from two sources S_1 and S_2 must be coherent.

3.2.1 Experimental Arrangement

Thomas Young was the first to demonstrate experimentally the phenomenon of interference. His experimental arrangement is shown schematically in Fig. 3.2. He allowed sunlight to fall on a single slit S , and then through double slits S_1 and S_2 . The interference pattern was observed on a screen which is placed at a distance D from S_1 and S_2 . Light waves coming from S_1 and S_2 have the same wavelength and same or constant phase relationship. Such waves are called **coherent waves**. Light from slits S_1 and S_2 interfere at different points resulting in constructive and destructive interference. Thus, on the screen alternate bright and dark bands of equal width called **interference fringes** are observed.

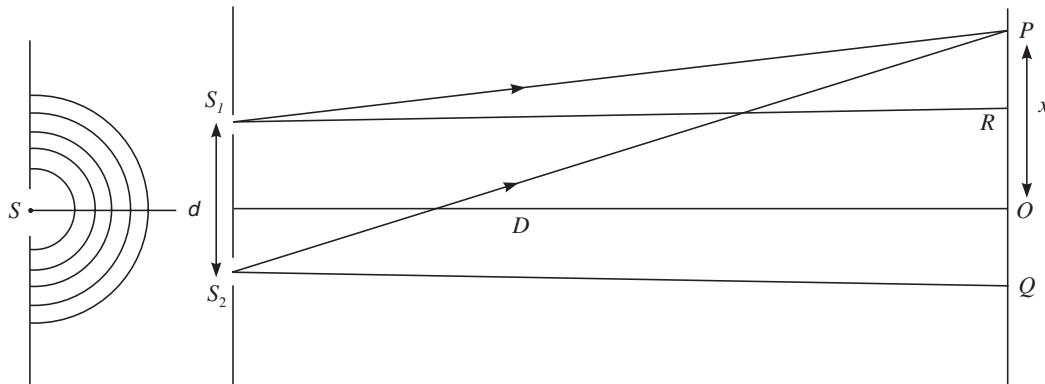


Fig. 3.2 Experimental arrangement for Young's experiment.

3.2.2 Expression for Bandwidth β

Let the source S be a monochromatic light source emitting light of wavelength λ . Let P be a point on the screen at a distance x from O . The point P will be bright or dark depending on the path difference between the rays from S_1 and S_2 reaching the point P . From Fig. 3.2, we get

$$\text{Path difference} = S_2P - S_1P$$

From triangles ΔS_1PR and ΔS_2PQ ,

$$S_1P^2 = D^2 + \left(x - \frac{d}{2}\right)^2$$

and,

$$S_2 P^2 = D^2 + \left(x + \frac{d}{2} \right)^2$$

$$S_2 P^2 - S_1 P^2 = \left(x + \frac{d}{2} \right)^2 - \left(x - \frac{d}{2} \right)^2 \quad (3.16)$$

$$(S_2 P + S_1 P)(S_2 P - S_1 P) = 2xd$$

Since

$$(S_2 P + S_1 P) \equiv 2D.$$

$$S_2 P - S_1 P = \frac{xd}{D} \quad (3.17)$$

For the point P to be bright, the path difference must be an integral multiple of λ . Hence, for the n th bright band,

$$\frac{x_n d}{D} = n\lambda \quad (3.18)$$

$$x_n = \frac{n\lambda D}{d} \quad (3.19)$$

For the $(n+1)$ th bright band,

$$x_{n+1} = \frac{(n+1)\lambda D}{d} \quad (3.20)$$

The distance between the centres of successive bright bands is then given by,

$$x_{n+1} - x_n = \frac{D\lambda}{d} \quad (3.21)$$

Similarly, for the n th dark band, path difference is given by

$$\frac{x_n d}{D} = (2n-1)\frac{\lambda}{2}$$

For the $(n+1)$ th dark band,

$$\frac{x_{n+1} d}{D} = [2(n+1)-1]\frac{\lambda}{2} = (2n+1)\frac{\lambda}{2}$$

The distance between successive dark bands is given by

$$x_{n+1} - x_n = \frac{D}{d} [(2n+1) - (2n-1)]\frac{\lambda}{2} = \frac{D\lambda}{d} \quad (3.22)$$

which is same as that of the bright bands. The term **band width** β is defined as the distance between the centres of two successive bright or dark bands

$$\beta = \frac{D\lambda}{d} \quad (3.23)$$

The band width is:

- (i) Directly proportional to the distance D between the slits and the screen.
- (ii) Directly proportional to the wavelength of the light used.
- (iii) Inversely proportional to the separation between the slits.

3.2.3 Conditions for Sustained Interference Pattern

To obtain a stable interference pattern certain conditions must be satisfied:

- (i) The two sources must be monochromatic, emitting light of same wavelength.
- (ii) The two sources must be coherent. That is, the waves should have equal amplitude, equal frequency and same or constant phase difference (Section 3.4).
- (iii) Separation between the two sources must be as small as possible.
- (iv) The two sources must be very narrow.

3.3 WHITE LIGHT FRINGES

When white light is used, as the source is polychromatic, the interference fringes will be coloured. White light consists of electromagnetic radiations having wavelengths ranging from violet (~ 400 nm) to red (750 nm). At the centre, all the wavelengths meet in phase and the central band is white. As we move towards the sides, the path difference gradually increases from zero. As fringe width is directly proportional to λ , the fringe width of red will be approximately double that for violet. At any point P , we may have maximum intensity for a number of wavelengths for which path difference is equal to $n\lambda$ and minimum intensity for other wavelengths for which path difference is equal to $(2n - 1)\lambda/2$. Due to overlapping of the intensity of different wavelengths, the point P will appear coloured if it is close to O . As the point P is moved further away from the centre, the path difference increases. As a result a large number of wavelengths will produce maximum intensity and an equally large number will produce minimum intensity at P . Hence, the resultant colour at P will nearly be white. Thus, when white light is used to illuminate the slit, we obtain an interference pattern consisting of a central white fringe with a few coloured symmetric fringes on both sides, and then a general illumination.

3.4 COHERENCE

Coherence is an important property associated with the idea of interference. Coherence means that two or more electromagnetic waves are in a fixed and predictable phase relationship to each other. In general, the phase between two electromagnetic waves can vary from point to point (in space) or change from instant to instant (in time). There are, thus, two independent concepts of coherence namely **spatial coherence** and **temporal coherence**.

Temporal coherence

Temporal coherence refers to the correlation between the field at a point and the field at the same point at a later time. If the phase difference between the two fields is constant during the period covered by observations, the wave is said to have temporal coherence. If the phase difference changes many times, and in an irregular manner during the shortest observation period, the wave is said to be in-coherent. The degree to which correlation exists is known as the **amount of longitudinal coherence**.

Spatial coherence

Spatial coherence refers to the continuity and uniformity of a wave in a direction perpendicular to the direction of propagation. If the phase difference for any two fixed points in a plane normal to the wave propagation does not vary with time, then the wave exhibits spatial coherence. This is possible even when two beams are individually time coherent, as long as any phase change in one of the beams is accompanied by a simultaneous equal phase change in the other beam. This is possible only if the two beams are produced in the same part of the source.

Light waves come in the form of wave trains because light is produced during de-excitation of electrons in atoms. These wave trains are of finite length, containing only a limited number of waves. The wave train lasts for less than roughly 10^{-8} s. This concept of wave trains is in agreement with certain experimental observations. The length of the wave train Δl is called the **coherence length**. It is the product of the number of waves N contained in the wave train and their wavelength λ . It takes a wave train of length Δl , a certain length of time Δt to pass a given point. Hence,

$$\Delta t = \frac{\Delta l}{c}$$

where c is the velocity of light. The time Δt is called **coherence time**. Temporal coherence is a characteristic of a single beam of light, whereas spatial coherence concerns the relationship between two separate beams of light.

3.5 THICKNESS OF TRANSPARENT SHEET

In Young's experiment, if a thin transparent sheet is introduced in the path of one of the beams, it is observed that there is a shift in the position of fringes. This is experimentally observed when white light is used as the source. From the fringe shift, we can determine the thickness of the sheet used.

3.5.1 Expression for Thickness of Sheet

Let a thin sheet of thickness t and refractive index μ be introduced in the path of one of the beams as shown in Fig. 3.3. It is seen that the central fringe is shifted to position P . This implies that the path difference of rays reaching P from S_1 and S_2 is zero. In other words, rays from S_1 and S_2 reach the point P at the same time. Hence,

$$\frac{S_2 P}{c} = \frac{(S_1 P - t)}{c} + \frac{t}{v} \quad (3.24)$$

where c is the velocity of light in free space and v is the velocity within the sheet. Multiplying Eq. (3.24) by c , we get

$$S_2 P = (S_1 P - t) + \mu t \quad \text{since } \mu = \frac{c}{v}$$

$$S_2 P - S_1 P = (\mu - 1) t \quad (3.25)$$

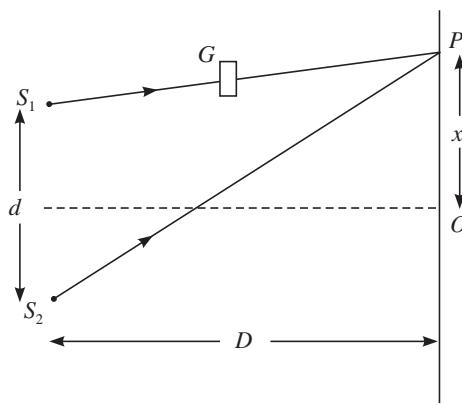


Fig. 3.3 Experimental arrangement to find the thickness of a thin sheet using Young's experiment.

From Eq. (3.17), we get

$$(S_2 P - S_1 P) = \frac{xd}{D} = (\mu - 1)t \quad (3.26)$$

Replacing white light with a monochromatic light and counting the number of bright fringes n between O and P , we get

$$\frac{xd}{D} = n\lambda \quad (3.27)$$

From Eqs. (3.26) and (3.27), we obtain

$$(\mu - 1)t = n\lambda$$

$$t = \frac{n\lambda}{\mu - 1} \quad (3.28)$$

Substituting the values of n , λ and μ the thickness t can be calculated.

3.6 METHODS TO PRODUCE INTERFERENCE PATTERN

Basically there are two methods to produce interference pattern. One method is by division of wavefront and the other method is by division of amplitude.

Division of wavefront

This method was discovered by Thomas Young. In the experiment two slits are illuminated by the light from a single source and wavefront is divided laterally into segments. It is similar to the interference pattern produced from two point sources vibrating in phase in a ripple water tank, for example, Fresnel's biprism, Lloyd's mirror, etc.

Division of amplitude

In this procedure a beam is divided into two or more reflecting surfaces and the reflected beams are allowed to interfere. The two resulting wavefronts maintain the original width, but have reduced amplitudes. Therefore, the intensity of the incident wavefront is divided. For example, Newton's rings (Section 3.9), Michelson's interferometer (Section 3.11), etc.

Optical path length

While discussing the division of amplitude cases, we require the concept of optical path length. If the velocity of light in a medium is v , the time taken to cover a distance t in the medium is $\frac{t}{v}$. The distance travelled by light in vacuum during this time is $\frac{ct}{v} = \mu t$ which is known as the **optical path length** in the medium of refractive index μ . That is,

$$t_{op} = \mu t \quad (3.29)$$

3.7 INTERFERENCE IN THIN FILMS

We are familiar with beautiful colours on an oil slick or by cracks in a piece of glass. This is due to the interference phenomenon involving multiple reflections. Here the method involved is interference by division of amplitude. The light reflected from the upper and lower surfaces of a thin film interferes and interference patterns are produced. The transmitted rays emerging from the lower side of the film can also interfere and produce interference patterns.

3.7.1 Reflected Light

Let us consider a thin film of thickness t and refractive index μ . Let light from a monochromatic source be incident on the surface of the film. Reflections take place from the upper and the lower surfaces of the film (Fig. 3.4). The reflected rays emerging from A and C have a common path up to A and equal path after CF .

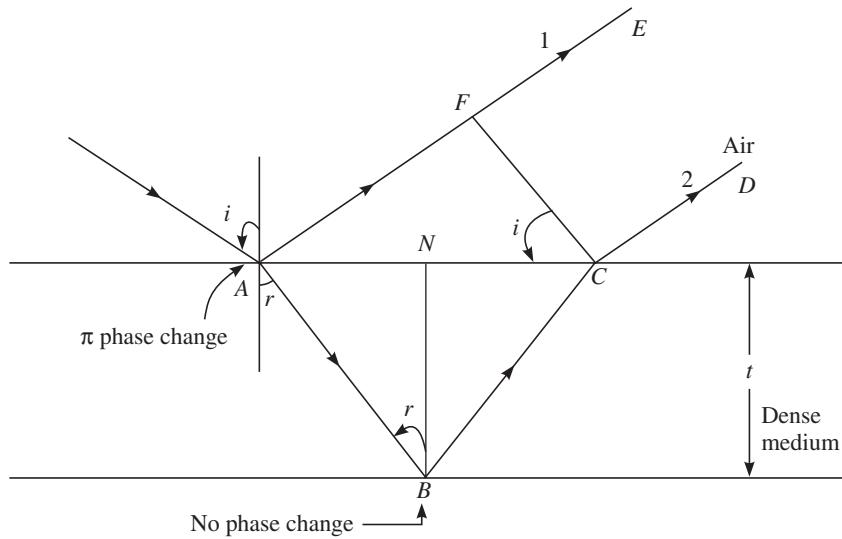


Fig. 3.4 Interference in thin films due to reflected light.

Hence, optical path difference between the two beams = $(AB + BC)$ in film – AF in air.
Since $AB = BC$,

$$\text{Optical path difference} = 2\mu AB - AF \quad (3.30)$$

In ΔANB ,

$$AB = \frac{t}{\cos r} \quad \text{and} \quad AN = t \tan r \quad (3.31)$$

In ΔAFC ,

$$\begin{aligned} AF &= AC \sin i = 2AN \sin i \\ AF &= 2t \tan r \sin i \end{aligned} \quad (3.32)$$

By Snell's law, $\sin i = \mu \sin r$. Hence,

$$AF = 2t \tan r \mu \sin r = 2\mu t \frac{\sin^2 r}{\cos r} \quad (3.33)$$

$$\begin{aligned} \text{Optical path difference} &= \frac{2\mu t}{\cos r} - \frac{2\mu t \sin^2 r}{\cos r} = \frac{2\mu t}{\cos r} (1 - \sin^2 r) \\ &= 2\mu t \cos r \end{aligned} \quad (3.34)$$

Whenever light is reflected from the surface of an optically denser medium, like the air-film interface, an additional phase change of π or a path difference of $\lambda/2$ for the reflected ray occurs. Therefore,

$$\text{The actual path difference} = 2\mu t \cos r + \frac{\lambda}{2}$$

(i) For constructive interference, we obtain

$$2\mu t \cos r + \frac{\lambda}{2} = n\lambda$$

$$2\mu t \cos r = (2n - 1) \frac{\lambda}{2}, \quad n = 1, 2, 3, \dots \quad (3.35)$$

(ii) For destructive interference, we obtain

$$\begin{aligned} 2\mu t \cos r + \frac{\lambda}{2} &= (2n + 1) \frac{\lambda}{2} \\ 2\mu t \cos r &= n\lambda, \quad n = 0, 1, 2, \dots \end{aligned} \quad (3.36)$$

(iii) If thickness of the film is very small ($t \sim 0$), then destructive interference occurs due to the additional path difference of $\lambda/2$ and the film appears dark. If t is very large, almost all the colours will undergo constructive interference, and a white light is produced.

Equation (3.36) is known as **cosine law**. Cosine law is a fundamental result in the study of various experiments like wedge shaped films (Section 3.8), Michelson's Interferometer (Section 3.11), etc.

Colour in thin films

When sunlight falls on a soap bubble, different colours are seen. This can be explained on the basis of Eqs. (3.35) and (3.36). The portion of the film which satisfies the condition for destructive interference ($2\mu t \cos r = n\lambda$) will be dark for that value of λ . For instance, if red satisfies this condition, then the spectrum will be devoid of red. A combination of other colours (VIBGYO) will be seen at that point. Also, in the case of oil film, where t and r vary with time, different colours are seen.

3.7.2 Transmitted Light

In the transmitted case, the two beams (Fig. 3.5) BE and DF interfere to produce interference pattern. The optical path difference between the two beams will be $(BC + CD)$ in film minus BH in air.

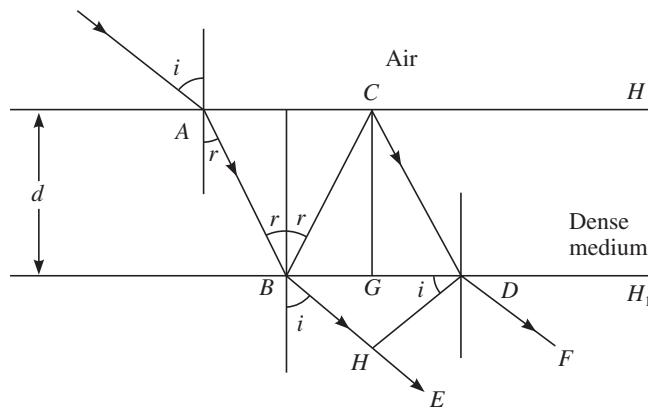


Fig. 3.5 Interference in thin films due to transmitted light.

Since $BC = CD$,

$$\begin{aligned}\text{Optical path difference} &= 2\mu BC - BH \\ &= 2\mu t \cos r\end{aligned}\quad (3.37)$$

Since the transmitted rays are from film-air interface, there is no additional path difference of $\lambda/2$. Hence, condition for constructive interference is given by

$$2\mu t \cos r = n\lambda; \quad n = 0, 1, 2, \dots \quad (3.38)$$

Condition for destructive interference is

$$2\mu t \cos r = (2n - 1) \frac{\lambda}{2}, \quad n = 1, 2, 3, \dots \quad (3.39)$$

From a comparison of the conditions for fringes in the reflected and transmitted systems, one can easily say that both are complementary. That is, for the same path difference a bright band in the reflected system will be a dark band in the transmitted system and vice versa.

3.8 WEDGE SHAPED FILMS

A wedge shaped film is constructed using two glass plates with one end placed in contact and the other end separated by a spacer of thickness t . The spacer can be a piece of paper a thin wire, etc. so that a thin film of varying thickness is formed between the two glass plates. The interference takes place between the light reflected from the upper and lower surfaces of the film (Fig. 3.6). The pattern obtained consists of straight alternate bright and dark bands of equal thickness. Let x_n represents position of n th dark band from 0.

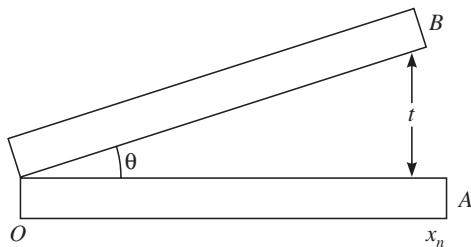


Fig. 3.6 Air wedge.

From Fig. 3.6, we get

$$\tan \theta = \frac{t}{x_n}$$

When θ is small,

$$\theta = \frac{t}{x_n} \quad \text{or} \quad t = \theta x_n$$

Substituting the value of t in Eq. (3.36) which is the condition for darkness in thin film interference, we get

$$2\mu \theta x_n = n\lambda \quad \text{or} \quad x_n = \frac{n\lambda}{2\mu\theta} \quad (3.40)$$

Let x_{n+m} represents the position of $(n + m)$ th dark fringe

$$x_{n+m} = \frac{(n+m)\lambda}{2\mu\theta} \quad (3.41)$$

$$x_{n+m} - x_n = \frac{m\lambda}{2\mu\theta} \quad (3.42)$$

What we measure in the experiment is this difference. If n and $(n + m)$ represent adjacent fringes then the above difference is band width β .

$$\beta = \frac{\lambda}{2\mu\theta}, \quad m = 1$$

For air film, $\mu = 1$

$$\beta = \frac{\lambda}{2\theta} \quad (3.43)$$

3.8.1 The Diameter of a Thin Wire

Let a wire of diameter d is used as the spacer in air wedge experiment. Using the same experimental arrangement, we can observe bright and dark fringes through a travelling microscope. Let x be the distance of the n th dark fringe from the edge. Then, $\theta = \frac{d}{x}$. Substituting this value of θ in Eq. (3.43), we get

$$\begin{aligned} \beta &= \frac{\lambda x}{2d} \\ d &= \frac{\lambda x}{2\beta} \end{aligned} \quad (3.44)$$

Band width β is measured experimentally and the diameter of the wire is determined.

3.8.2 Testing the Planeness of Surfaces

Air wedge experiment can be used to test the optical planeness of surfaces. The surface to be tested is used as one of the plates in air wedge experiment. The other will be a standard plane surface. If the surface is optically plane, then straight fringes of equal thickness are observed. If the fringes are non-uniform and distorted, the given surface is not optically plane. A surface is said to be *optically plane* if it is flat up to 1/10th of the wavelength of the light used. With the air wedge experiment, flatness of a surface up to 1/10th of the wavelength of the light used can be determined.

3.9 NEWTON'S RINGS BY REFLECTED LIGHT

As pointed out earlier, this is one of the applications of producing interference pattern by division of amplitude.

3.9.1 Experimental Arrangement

A plano convex lens L_2 is placed on the surface of a plane glass plate. A thin air film is formed between the curved surface of the lens and the surface of the glass plate. Light reflected from the curved surface of the lens and from the surface of the glass plate interfere to produce interference pattern (refer Fig. 3.7). The thickness of the film is zero at the point of contact (at O) and increases radially outwards. The thickness of the air film will be constant over a circle and the pattern consists of concentric bright and dark rings as shown in Fig. 3.8(a). These rings are referred to as **Newton's rings**. From the interference pattern, it may be noted that the centre spot is dark. The thickness of the film is zero at O . Therefore, only a path difference of $\lambda/2$ results due to the reflection from the surface of the glass plate G and this favours destructive interference. That is why the centre spot is dark in Newton's rings. From there, as the thickness of the film increases, depending on the condition, alternate bright and dark rings are formed.

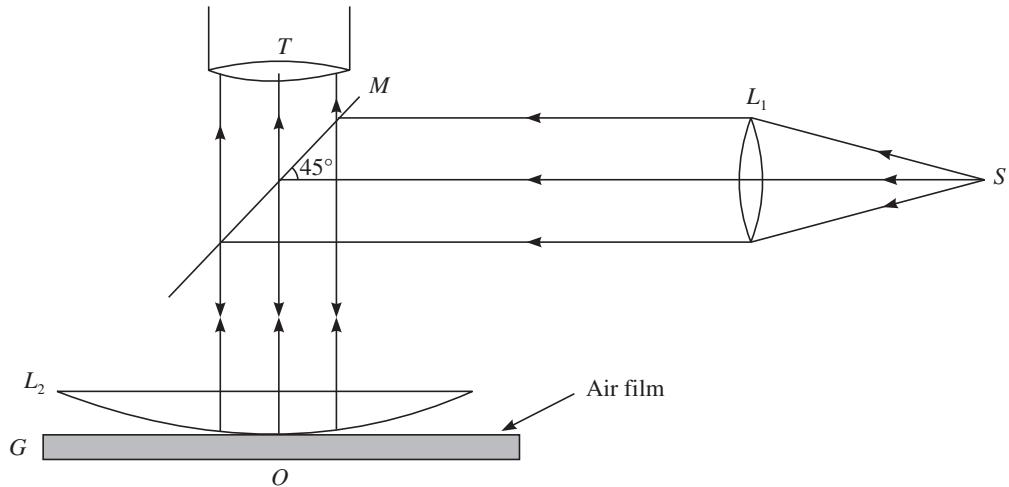


Fig. 3.7 Experimental arrangement for Newton's rings.

Radius of the n th ring

Let r_n be the radius of the n th dark ring. [Fig. 3.8(b)] R , the radius of curvature of the plano-convex lens, and λ the wave length of light used. The condition for minimum intensity in thin film interference as in Eq. (3.36) is given as

$$2\mu t \cos r = n\lambda$$

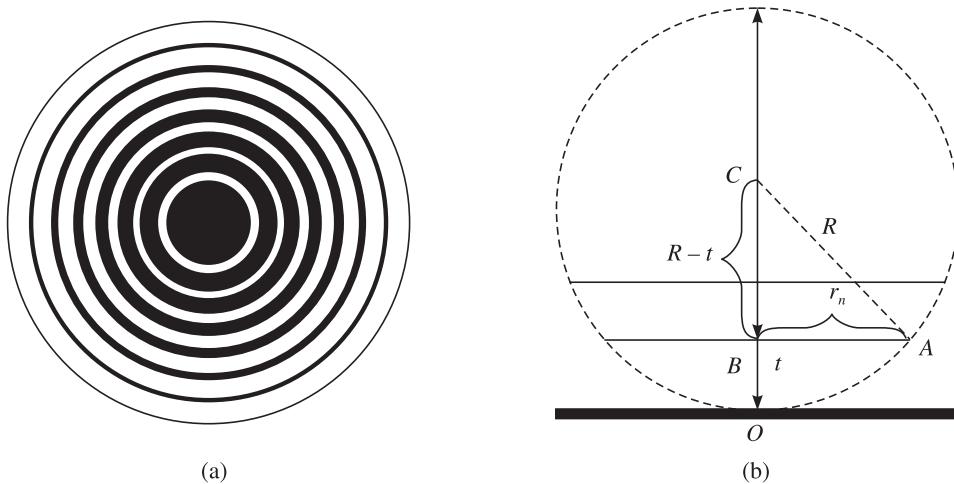


Fig. 3.8 Newton's rings in reflected light (a) Interference pattern (b) r_n represents the radius of the n^{th} dark ring where the thickness of the film is t .

Since light is incident on the film almost normally $\cos r = 1$ and for air film $\mu = 1$. Hence,

$$2t = n\lambda \quad (3.45)$$

From Fig. 3.8(b), considering ΔABC

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

$$R^2 = r_n^2 + (R - t)^2$$

Neglecting t^2 , we get

$$2Rt = r_n^2 \quad \text{or} \quad t = \frac{r_n^2}{2R} \quad (3.46)$$

Substituting this value of t in Eq. (3.45), we get

$$r_n = \sqrt{Rn\lambda}, \quad n = 0, 1, 2, \dots \quad (3.47)$$

Similarly for n^{th} bright ring,

$$r_n = \sqrt{R(2n-1) \frac{\lambda}{2}}, \quad n = 1, 2, 3, \dots \quad (3.48)$$

Here, n is the order of the ring. If dust particles are present in the space between the glass plate and lens, we will not be able to determine n correctly. Therefore, it is the practice to avoid n from the above expression. If D_n and D_{n+k} are the diameters of the n^{th} and $(n+k)^{\text{th}}$ dark rings, respectively, then

$$D_n^2 = 4Rn\lambda \quad (3.49)$$

and

$$D_{n+k}^2 = 4R(n+k)\lambda \quad (3.50)$$

$$\lambda = \frac{D_{n+k}^2 - D_n^2}{4Rk} \quad (3.51)$$

Newton's rings are observed directly through the travelling microscope T (Fig. 3.7) and measurements of the diameter of n th and $(n + k)$ th dark rings are made. Substituting the values of the diameter in Eq. (3.51), the wavelength of the light used can be calculated.

Newton's rings with white light

With white light, a few coloured rings appear near the point of contact. Ring diameters depend on wavelength (Eq. 3.49). Due to overlapping, higher order rings will not be observed as we move outward.

3.10 NEWTON'S RINGS WITH LIQUID FILM

If a liquid of refractive index μ is introduced between the lens and the glass plate, the optical path difference becomes $2\mu t$. Let d_{n+k} and d_n be the diameters of $(n + k)$ th and n th dark rings, respectively. Then,

$$\begin{aligned} d_n^2 &= \frac{4nR\lambda}{\mu}; \quad d_{n+k}^2 = \frac{4(n+k)R\lambda}{\mu} \\ \lambda &= \mu \frac{(d_{n+k}^2 - d_n^2)}{4Rk} \end{aligned} \quad (3.52)$$

The experiment is repeated with air film in between. Let D_{n+k} and D_n represents the diameter of the corresponding rings. Then,

$$\lambda = \frac{D_{n+k}^2 - D_n^2}{4Rk} \quad (3.53)$$

Dividing Eq. (3.52) by Eq. (3.53), we get

$$\mu = \frac{D_{n+k}^2 - D_n^2}{d_{n+k}^2 - d_n^2} \quad (3.54)$$

It can also be seen that

$$\mu = \frac{r_m^2}{r_m'^2} = \frac{D_m^2}{d_m^2} \quad (3.55)$$

where r_m and r'_m represent the radius of the same bright or dark ring with and without the liquid film. Since μ is always greater than 1,

$$r_m > r'_m \quad (3.56)$$

Hence, we can conclude that the rings contract with the introduction of a liquid. Newton's rings technique can be used to check if two surfaces such as lenses have the same curvature. It can also be used to check whether or not a lens is suitable for use in optical instruments intended

for quality work. Another practical application is the determination of the optical flatness of a glass plate.

3.10.1 Ring System in Transmitted Light

Here the rings are exactly complementary to the reflected ring system, so that the center spot is bright. As the amplitudes of the interfering rays are not equal, the transmitted rings are less distinct than the rings formed by reflected light.

3.11 MICHELSON INTERFEROMETER

3.11.1 Principle

Michelson interferometer is the best example of interference pattern by division of amplitude. The amplitude of light beam from a source is divided into two parts of equal intensities by partial reflection and transmission. These beams are then sent in two directions at right angles and are then brought together by the use of two highly polished plane mirrors to produce interference fringes.

3.11.2 Experimental Arrangement

A schematic diagram of Michelson's interferometer is given in Fig. 3.9. Light from source S falls on a glass plate G_1 which has a reflective coating on the back. The coating reflects half of the beam to the plane mirror M_1 . The transmitted beam of equal intensity travels to plane mirror M_2 . The beams reflected by M_1 and M_2 are recombined again after passing through G_1 . The reflected beam goes through the plate G_1 three times, whereas the other beam goes through G_1 only once. Hence, to make the optical paths of the two beams equal, a compensating plate G_2 is placed in the path of the transmitted beam. Plane mirrors M_1 and M_2 have arrangements to make them exactly perpendicular to each other.

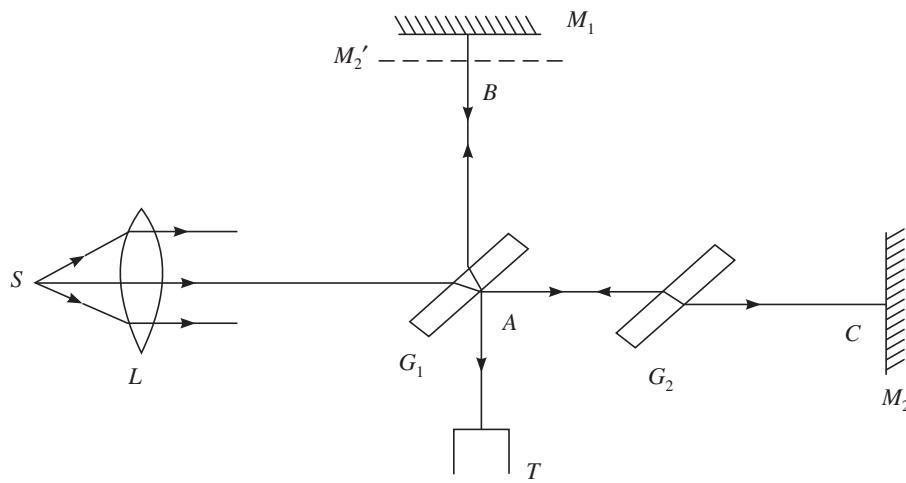


Fig. 3.9 Michelson interferometer.

3.11.3 Working

The two beams arrive at the telescope T and produce interference fringes. Mirror M_1 can be moved along the axis exactly parallel to itself. M'_2 is the virtual image of M_2 formed by reflection in G_1 . Depending on the position of the mirrors, the image M'_2 may be in front of or exactly coincident with or behind mirror M_1 . If the mirror M_1 is moved a distance d , the light passes through the distance twice and the path length changes by $2d$.

The light that comes from M_2 and goes to T is reflected at A in an air to glass boundary and there will be an additional phase change of π . The situation for the other beam is slightly more complicated. A phase change occurs for that too since the reflection is at glass-metallic interface in G_1 which is not exactly π . Neglecting the slight difference (valid for very thin coating) path difference is $2d$. This means that at the coincidence position there will be a minimum, and the center of the field will be dark. If the mirror M_1 is moved through a distance $\lambda/4$, the path length changes by $\lambda/2$ and we have a maximum. Thus, constructive interference occurs when

$$2d = m\lambda, \quad m = 0, 1, 2, \dots$$

If we look obliquely at the interferometer, in such a way that the line of sight makes an angle θ with the axis, the path difference is nearly $2d \cos \theta$. Hence, the condition for maximum is given by

$$2d \cos \theta = m\lambda \quad (3.57)$$

For a given mirror separation d , θ is constant. The fringes appear in the form of concentric circles around the axis. These are fringes of **equal inclination**. If the mirrors are tilted, the fringes appear straight. Actually they are sections of large hyperbolas called **localized fringes**.

White light fringes

When a source of white light is used, fringes will be observed only when the path difference do not exceed a few wavelengths. The mirrors are titled slightly for observing these fringes. There will be a central dark fringe followed by a few coloured fringes on either side. When we move M_1 slowly, more fringes will come to the view. Normally, in a white light system we expect central fringe to be white, since the two beams should be in phase with each other for any wavelength. But, when observed in Michelson interferometer, the central fringe is dark. This is because the back surface of G_1 is unsilvered and the ray towards AB undergoes an internal reflection in the plate G_1 , ray towards AC undergoes an external reflection, with a consequent change of phase. If the back surface of G_1 is silvered, the conditions are different and the central fringe will be white.

3.11.4 Applications

Of all the interferometers, Michelson interferometer is the most versatile one. It makes it possible to bring two optical planes into coincidence and to move them virtually through one another. It can be used with an extended source. Compared to a set of double slits, it gives much brighter fringes.

Determination of wavelength

The monochromatic light whose wavelength is to be determined is allowed to fall on the plate G_1 . Mirrors M_1 and M_2 are made exactly perpendicular so that circular fringes are seen. With M_2 fixed, the position of M_1 is adjusted to obtain a bright spot at the centre of the field of view. If M_1 is further moved a distance equal to $\lambda/2$, $2d$ changes by λ and we will again get a bright spot at the centre. The order of the spot now is $(n + 1)$. Thus, one fringe crosses the field of view each time when M_1 is moved a distance equal to $\lambda/2$. If N fringes cross the field of view and x is the distance through which M_1 is moved, then λ can be calculated from the equation,

$$x = \frac{N\lambda}{2} \quad \text{or} \quad \lambda = \frac{2x}{N} \quad (3.58)$$

Determination of small difference in wavelengths

Let λ_1 and λ_2 be two close wavelengths emitted by a source, and let $\lambda_1 > \lambda_2$. Adjustment is done to obtain bright distinct circular fringes by adjusting mirror M_1 . At this condition, bright fringe due to λ_1 coincides with the bright fringe due to λ_2 . When the mirror M_1 is moved further, the visibility of the fringes decreases and tends to zero. The field now appears uniformly illuminated. At this condition, the dark fringes of one set fall on the bright fringes of the other resulting in maximum indistinctness. On further moving M_1 , the visibility increases and again a position comes when the fringes are well-defined and very distinct. Let x be the distance moved by the mirror for two successive positions of maximum distinctness or indistinctness, then the path difference is $2x$. During this movement, let n be the number of fringes moved of longer wavelength λ_1 and $(n + 1)$ that of λ_2 . Then, we have

$$2x = n\lambda_1 = (n + 1)\lambda_2 \quad (3.59)$$

$$n = \frac{\lambda_2}{\lambda_1 - \lambda_2} \quad (3.59a)$$

With this value of n in Eq. (3.59), we obtain

$$\begin{aligned} 2x &= \frac{\lambda_2}{\lambda_1 - \lambda_2} \cdot \lambda_1 \\ \lambda_1 - \lambda_2 &= \frac{\lambda_1 \lambda_2}{2x} \end{aligned} \quad (3.60)$$

As λ_1 and λ_2 are close, $\lambda_1 \lambda_2 = \lambda^2$ where, $\lambda = \frac{(\lambda_1 + \lambda_2)}{2}$. Hence,

$$\Delta\lambda = \lambda_1 - \lambda_2 = \frac{\lambda^2}{2x} \quad (3.61)$$

Using Eq. (3.61), the difference in wavelength can be calculated.

Determination of refractive index of thin film

The interferometer is adjusted to produce white light fringes. The telescope is focused with its cross wire at the central dark fringe. When a film of thickness t and refractive index μ is

introduced in the path of one of the interfering beams, the path difference is $2(\mu - 1)t$ and there is shift in the central fringe. Now, the mirror M_1 is moved so that the central fringe is again brought to its initial position. If x is the distance through which the mirror is moved, then

$$\begin{aligned} 2x &= 2(\mu - 1)t \\ x &= (\mu - 1)t \end{aligned} \quad (3.62)$$

Thus, by measuring x and t , μ may be calculated or if μ is known, we can calculate t .

Standardisation of the meter

Michelson's interferometer is used to determine the length of the standard meter in terms of the wavelength of the orange line in the spectrum of krypton—86. Repeated measurements of the old standard meter (the one kept in Paris) were made (1960) to establish one meter as being 1650763.73 wavelengths of this light. In 1983, the meter was again redefined in terms of the speed of light.

3.12 INTERFERENCE FILTER

Interference filter is a device which transmits nearly a monochromatic beam of light with a very narrow band width. It consists of a thin transparent dielectric such as magnesium fluoride (MgF_2) or cryolite (Na_3AlF_6) sandwiched between two plane glass plates. Semi-transparent silver films are deposited on the faces of the glass plates as shown in Fig. 3.10. When light falls normally on such an arrangement, multiple reflections take place within the dielectric film. The interference maxima of different orders are formed in the transmitted light only for those wavelengths which satisfy the relation.

$$2\mu t = n\lambda,$$

where n is the integer, μ is the refractive index of the dielectric and t its thickness. The lower glass plate is often called the **substrate** and the upper one the **cover plate**.

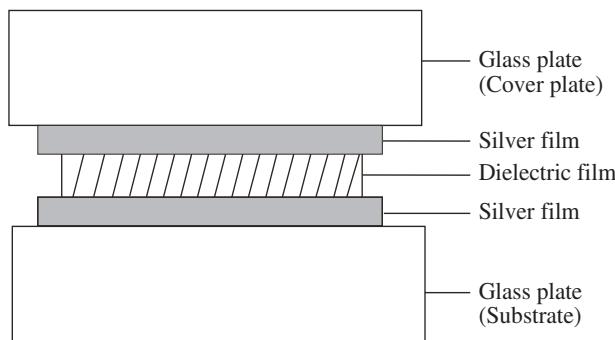


Fig. 3.10 The interference filter.

The optical thickness of the dielectric is μt . For $\mu t = 0.5 \times 10^{-6}$ m, the transmitted wavelengths corresponding to maxima are given by $\lambda = 2\mu t/n$, $n = 1, 2, 3, \dots$. For $\mu t = 0.5 \times 10^{-6}$ m gives λ values, 10000 Å, 5000 Å, 3333 Å, Of these, only 5000 Å lies in the visible region.

By varying the thickness of the dielectric, interference filter can be worked at any desired wavelength. The sharpness of the transmitted spectrum is determined by the reflectivity of the metallic surfaces. Larger the reflectivity, narrower is the transmitted spectrum. However, it is not possible to increase the thickness of the metallic film beyond a limit, since this would reduce the intensity of the transmitted light. To overcome this difficulty, metallic films are replaced by dielectric structures. Sometimes, multilayer structures of higher and lower refractive indices are also used. Such interference filters are capable of transmitting a band of wavelength as small as 11 Å or even with peak at any wavelength within the visible region.

3.13 ANTIREFLECTION COATINGS

One of the important applications of thin film interference lies in reducing the reflectivity of lens surfaces. In many optical instruments like telescope, range finders etc., the incident light has to undergo reflections from many surfaces. When the light reflects back from each surface, there occurs considerable loss in the intensity of the transmitted light. This loss can be reduced by coating the lens surface with a suitable transparent dielectric material whose refractive index is intermediate between air and glass.

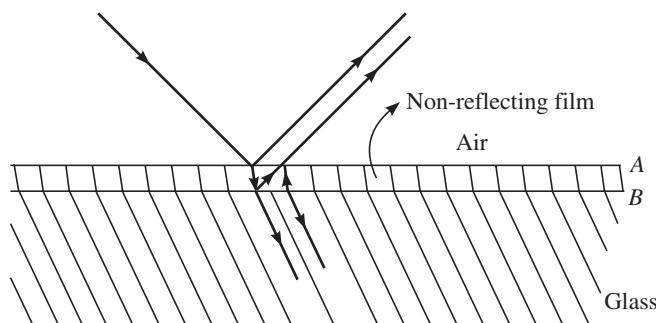


Fig. 3.11 Antireflection coating.

The incident light is reflected from the upper and lower surfaces *A* and *B* of the film. (refer Fig. 3.11). Since both the rays are reflected at a rare medium-denser medium interface, the same phase change π occurs in both cases. If the thickness of the film is such that both the reflected rays are in opposite phase, they will cancel each other due to destructive interference. For destructive interference, we obtain

$$2\mu t = \frac{\lambda}{2}$$

Then, $t = \lambda/4\mu$ or any odd integral multiple of $\lambda/4\mu$. The thickness is suitable only for a particular wavelength. So the wavelength is selected from the yellow-green region of the visible spectrum where the eye is most sensitive. The intensity of the transmitted light is increased as there is no loss due to reflection. The coating materials used are magnesium fluoride, cryolite etc. They are usually deposited by vacuum evaporation technique.

SOLVED EXAMPLES

Example 3.1 In Young's double slit experiment the separation of the slits is 2.2 mm and the fringe spacing is 0.51 mm at a distance of 2 m from the slits. Calculate the wavelength of light.

Solution: $\beta = 0.51 \text{ mm} = 0.051 \text{ cm}$

$$d = 2.2 \text{ mm} = 0.22 \text{ cm}$$

$$D = 2 \text{ m} = 200 \text{ cm}$$

$$\beta = \frac{D\lambda}{d} \quad \text{or} \quad \lambda = \frac{\beta d}{D}$$

$$\lambda = \frac{0.051 \text{ cm} \times 0.22 \text{ cm}}{200 \text{ cm}} = 5610 \times 10^{-8} \text{ cm} = 5610 \text{ \AA}$$

Example 3.2 A light source emits lights of two wavelengths $\lambda_1 = 4250 \text{ \AA}$ and $\lambda_2 = 5050 \text{ \AA}$. The source is used in a double slit interference experiment. The distance between the sources and the screen is 1.5 m and the distance between the slits is 0.025 mm. Calculate the separation between the third bright fringe due to these two wavelengths.

Solution: From Eq. (3.19), we obtain

$$x_3 = \frac{n\lambda_1 D}{d} \quad \text{and} \quad x'_3 = \frac{n\lambda_2 D}{d}$$

$$x'_3 - x_3 = \frac{nD}{d} (\lambda_2 - \lambda_1)$$

$$= \frac{3 \times 1.5 \text{ m}}{(0.025 \times 10^{-3} \text{ m})} (5050 \text{ \AA} - 4250 \text{ \AA})$$

$$= 1.44 \text{ cm}$$

Example 3.3 When a thin sheet of glass $3.9 \times 10^{-4} \text{ cm}$ thick is placed in the path of one of the interfering beams in Young's experiment, it is found that the central bright fringe is shifted to the position previously occupied by the fourth fringe. Find the refractive index of the sheet of glass. Wavelength of light used is $5.5 \times 10^{-5} \text{ cm}$.

Solution: From Eq. (3.28), we have

$$\mu = \frac{n\lambda}{t} + 1$$

where n is the number of fringes shifted

$$\mu = \frac{4(5.5 \times 10^{-5} \text{ cm})}{3.9 \times 10^{-4} \text{ cm}} + 1 = 1.5641$$

72 Engineering Physics

Example 3.4 A monochromatic light of wavelength 5893 Å is incident normally on a soap film of refractive index 1.42. What is the least thickness of the film that will appear (i) bright and (ii) dark by reflection.

Solution: (i) Condition for brightness is

$$2\mu t \cos r = (2n - 1) \frac{\lambda}{2}$$

For normal incidence, $\cos r = 1$ and for least thickness of the film, $n = 1$

$$t = \frac{\lambda}{4\mu} = \frac{5893 \text{ \AA}}{4 \times 1.42} = 1037.5 \text{ \AA}$$

(ii) Condition for darkness is

$$2\mu t \cos r = n\lambda$$

$$t = \frac{\lambda}{2\mu} = \frac{5893 \text{ \AA}}{2 \times 1.42} = 2075 \text{ \AA}$$

Example 3.5 Interferene fringes are formed in a thin air wedge using sodium light of wavelength 5893 Å. When observed normally, 10 fringes are found in a distance of 1 cm. Calculate the angle of the wedge.

Solution: Band width $\beta = \frac{1 \text{ cm}}{10} = 0.1 \text{ cm}$

From Eq. (3.43), the angle of the wedge θ is given by

$$\theta = \frac{\lambda}{2\beta} = \frac{5893 \times 10^{-8} \text{ cm}}{2 \times 0.1 \text{ cm}} = 2.946 \times 10^{-4} \text{ rad}$$

Example 3.6 An air wedge is formed using two optically plane glass strips of length 10 cm. A spacer of thickness 0.010 mm is introduced at one end to form an air film. If the light used is of wavelength 5900 Å find the separation between consecutive bright fringes.

Solution: From Eq. (3.43), band width $\beta = \frac{\lambda}{2\theta}$

$$\theta = \frac{\text{Spacer thickness}}{\text{Wedge length}} = \frac{0.001 \text{ cm}}{10 \text{ cm}} = 10^{-4}$$

$$\beta = \frac{5900 \times 10^{-8} \text{ cm}}{2 \times 10^{-4}} = 0.295 \text{ cm}$$

Example 3.7 In a Newton's ring experiment, the diameter of 4th and 12th dark rings are 0.4 cm and 0.7 cm, respectively. Deduce the diameter of 20th dark ring.

Solution: We know that

$$D_{n+k}^2 - D_n^2 = 4kR\lambda$$

∴ For diameters of 4th and 12th dark rings, we have

$$\begin{aligned} D_{12}^2 - D_4^2 &= 32R\lambda \quad \text{and} \quad D_{20}^2 - D_{12}^2 = 32R\lambda \\ D_{20}^2 - D_{12}^2 &= D_{12}^2 - D_4^2 \quad \text{or} \quad D_{20}^2 = 2D_{12}^2 - D_4^2 \\ D_{20}^2 &= 2(0.7\text{ cm})^2 - (0.4\text{ cm})^2 = 0.82\text{ cm}^2 \\ D_{20} &= 0.9055\text{ cm} \end{aligned}$$

Example 3.8 When a liquid is introduced between the lens and the plate in a Newton's ring arrangement, the diameter of the n th dark ring changes from 0.30 cm to 0.25 cm. Calculate the refractive index of the liquid.

Solution: With air film $D_n^2 = (0.30)^2 \text{ cm}^2$

With liquid film $d_n^2 = (0.25)^2 \text{ cm}^2$

$$\text{From Eq. (3.55)} \quad \mu = \frac{D_M^2}{d_m^2} = \left(\frac{0.30\text{ cm}}{0.25\text{ cm}} \right)^2 = 1.44$$

Example 3.9 When the movable mirror of Michelson's interferometer is shifted through a distance of 0.002945 cm, the shift of 100 fringes is observed. Calculate the wavelength of light in Å.

Solution: From Eq. (3.58), we get

$$\begin{aligned} \lambda &= \frac{2x}{N} = \frac{2(0.002945 \times 10^{-2} \text{ m})}{100} \\ &= 5890 \times 10^{-10} \text{ m} = 5890 \text{ Å} \end{aligned}$$

Example 3.10 Michelson's interferometer experiment is performed with D_1 and D_2 lines of wavelengths 5896 Å and 5890 Å. Find the distance through which the mirror has to be moved between two consecutive positions of maximum distinctness.

Solution: From Eq. (3.61), we obtain

$$\lambda_1 - \lambda_2 = \frac{\lambda^2}{2x}; \quad \lambda = \frac{\lambda_1 + \lambda_2}{2}$$

In the given case

$$\begin{aligned} \lambda &= \frac{5896 \text{ Å} + 5890 \text{ Å}}{2} = 5893 \text{ Å} \\ x &= \frac{\lambda^2}{2(\lambda_1 - \lambda_2)} = \frac{(5893 \text{ Å})^2}{2 \times 6 \text{ Å}} \\ &= 2.894 \times 10^{-2} \text{ cm} = 0.289 \text{ mm} \end{aligned}$$

REVIEW QUESTIONS

- 3.1** What are the conditions that should be satisfied to obtain interference fringes?
- 3.2** Why two independent sources of light cannot produce an interference pattern?
- 3.3** Define optical path.
- 3.4** Why is the centre of Newton's ring dark in the reflected system?
- 3.5** What happens when white light is used as the source in Young's experiment?
- 3.6** What are coherent sources? Distinguish between spatial coherence and temporal coherence.
- 3.7** Mention some of the applications of Michelson's interferometer.
- 3.8** Explain the conditions for obtaining circular fringes and straight line fringes using Michelson's interferometer.
- 3.9** How will you test the optical planeness of a glass plate using an air wedge?
- 3.10** What are interference filters?
- 3.11** Write a note on antireflection coatings.
- 3.12** For constructive interference between two waves of equal amplitude show that the intensity is four times than that of the individual wave.
- 3.13** Discuss some of the applications of interference.
- 3.14** What is cosine law? Explain.
- 3.15** The intensity of the transmitted Newton's rings system is less than that of the reflected system. Why?
- 3.16** A double slit is illuminated first by red light and then by violet light. Which colour gives the wider interference pattern beyond the double slit?
- 3.17** With examples, distinguish between interference by division of wavefront and by division of amplitude.
- 3.18** What is an interference filter? In what respects it differs from an ordinary filter?
- 3.19** What is an antireflection coating? What is it used for?
- 3.20** Derive an expression for bandwidth in Young's double slit experiment. Explain the effect of the introduction of a transparent material in one of the beams of double slit experiment.
- 3.21** Describe theoretically how the thickness of a thin sheet of transparent material is determined using Young's experiment.
- 3.22** Explain interference by thin films. Show that the colours exhibited by reflected and transmitted systems in thin film interference are complementary.
- 3.23** What are Newton's rings? Explain how the wavelength of light is determined using Newton's rings.
- 3.24** Describe with the necessary theory how the refractive index of a liquid is determined using Newton's rings arrangement.

- 3.25** Distinguish between interference by division of wavefront and by division of amplitude. Describe how Michelson's interferometer is used to determine the difference between two close wavelengths.
- 3.26** Explain with a diagram Michelson's interferometer. How is it used to find the thickness of a mica sheet?
- 3.27** Explain the formation of interference fringes using air wedge. How is it used to determine the wavelength of light?

PROBLEMS

1. Newton's rings are formed by reflection in the air film between a plane surface and a spherical surface of radius 50 cm. If the radius of the third bright ring is 0.09 cm and that of the twenty third 0.25 cm, what is the wavelength of the light used?
2. Interference fringes are produced by a thin wedge shaped film of plastic of refractive index 1.4. If the angle of the wedge is 20 seconds of arc and the distance between fringes is 0.25 cm, find the wavelength of the monochromatic light incident perpendicularly on the plastic.
3. With two slits spaced 0.2 mm apart, and a screen at a distance of 1m, the third bright fringe is found to be displaced 7.5 mm from the central fringe. Find the wave length of the light used.
4. Two slits are spaced 0.3 mm apart, and are placed 50 cm from a screen. What is the distance between the second and third dark line of the interference pattern when the slits are illuminated with light of 600 nm wavelength?
5. Light from a mercury arc lamp is passed through a filter that blocks everything except for one spectrum line in the green region of the spectrum. It then falls on two slits separated by 0.6 mm. In the resulting interference pattern on a screen 2.5 m away, adjacent bright fringes are separated by 2.27 mm. What is the wavelength?
6. Two rectangular pieces of plane glass are laid one upon the other on a table. A thin strip of paper is placed between them so that an air wedge is formed. This is illuminated by a beam of Na light at normal incidence ($\lambda = 589 \text{ nm}$). Interference fringes are formed, there being ten fringes per centimetre length of wedge measured normal to the edges in contact. Find the angle of the wedge.
7. The radius of curvature of the convex surface of a plano-convex lens is 120 cm. The lens is placed convex side down on a plane glass plate, and illuminated from above with red light of wavelength 650 nm. Find the diameter of the third bright ring in the interference pattern.
8. A source emits light of two wavelengths 6000 \AA and 4500 \AA . Newton's rings arrangement was set up using the source, and interference pattern is observed. It is found that n th dark ring due to 6000 \AA coincides with $(n + 1)$ th dark ring due to 4500 \AA . Radius of curvature of curved surface is 90 cm. Find the diameter of n th dark ring due to 6000 \AA .

76 Engineering Physics

9. Red light of wavelength 750 nm from a narrow slit is incident on a double slit. If the overall separation of 10 fringes on a screen 300 cm away is 4 cm, find the slit separation.
10. Two coherent sources are 0.20 mm apart and the fringes are observed on a screen 90 cm away. It is found that with a certain monochromatic source of light, the fourth bright fringe is situated at a distance of 12.8 mm from the central fringe. Calculate the wavelength of light.
11. A wedge shaped air film having an angle of 0.48×10^{-3} radians is illuminated by a monochromatic light and fringes are observed normally. Calculate the wavelength of the light used if the fringe width is 0.04 cm.
12. In Newton's rings arrangement, using a light of wavelength 546 nm, the radius of the n th and $(n + 20)$ th bright rings are found to be 0.162 cm and 0.368 cm, respectively. Calculate the radius of curvature of the lens.
13. A parallel beam of light of wavelength 58 nm is incident on a thin film of refractive index 1.5. Calculate the smallest thickness of the plate which would appear dark by reflection, if the angle of refraction into the plate is 60° .
14. In Young's experiment, a thin transparent sheet of thickness 6.3×10^{-4} cm is introduced in the path of one of the interfering beams. It is found that the central fringe is shifted to a position previously occupied by the sixth bright fringe. If $\lambda = 5460 \text{ \AA}$, find the refractive index of the sheet.
15. Using Newton's rings arrangement, the interference pattern is formed with a liquid film between the lens and the glass plate. Wavelength of light used is 589 nm. If the radius of curvature of the curved surface is 100 cm and the diameter of the 5th dark ring 0.3 cm, calculate the refractive index of the liquid.
16. When a micasheet of thickness 0.005 cm is introduced in front of the fixed mirror in Michelson's interferometer experiment, a shift is produced for white light straight fringes. In order to bring back the coloured fringes to the original position, the movable mirror is moved by 0.0025 cm. Calculate the refractive index of the sheet.
17. If the movable mirror of Michelson's interferometer is shifted through a distance of 0.0589 mm, 200 fringes cross the field of view. Calculate the wavelength of the light used.
18. A non-reflecting layer is formed with a coating of refractive index 1.334. What should be the thickness at wavelength 5500 \AA for zero reflection?
19. In Newton's rings experiment, the diameters of 4th and 10th dark rings are 0.4 cm and 0.6 cm, respectively. Deduce the diameter of 16th dark ring.
20. A shift of 100 fringes is observed when the movable mirror of Michelson's interferometer moves through a distance of 0.02395 mm. Calculate the wavelength of light used.
21. Two pieces of plane glass plates are placed together with a piece of paper between the two at one edge. Find the angle of wedge shaped air film in between them, if monochromatic light of wavelength 4800 \AA is incident on it and band width = 0.0555 cm.
22. Two optically plane glass plates of length 10 cm are placed one above the other. A thin foil of thickness 0.01 mm is introduced between the plates at one end to form an air wedge. If the light used has a wavelength 5900 \AA , find the separation between consecutive bright fringes.

23. Newton's rings are formed in reflected light of wavelength 6000 \AA with a liquid between the plane and curved surfaces. If the diameter of the sixth dark ring is 3.1 mm and the radius of curvature of the curved surface is 100 cm, calculate the refractive index of the liquid.
24. A fine metal foil separates one end of two pieces of optically flat glass. When light of wavelength 500 nm is incident normally, 28 dark lines are observed (with one at each end). Find how thick is the foil.
25. A soap bubble 250 nm thick is illuminated by white light. The index of refraction of the soap film is 1.36, find out which colours are seen in the reflected light, which colours appear strong in the reflected light and what colour does the soap film appear at normal incidence?

C H A P T E R

4

Diffraction

When light passes through a narrow circular aperture, it spreads out more than what could be accounted for by geometric considerations. This is one of the simplest examples of diffraction. Hence, diffraction can be defined as the failure of light to travel in straight lines. In this chapter we shall investigate the diffraction pattern in certain cases using the principle of wave motion.

4.1 THE DIFFRACTION PHENOMENON

When a small obstacle or an aperture is placed in the path of light, some light is found in the geometric shadow. Careful observation shows that if the width of the slit is not very large as compared to the wavelength of light used, then the intensity in the illuminated region is also not uniform. If the width of the slit is made smaller, larger amounts of energy reach the geometrical shadow. In other words, light does not propagate strictly in a straight line when it passes close to the edges of objects, and the patterns produced depend on the shape and size of the object. This phenomenon of bending of light around the edges of obstacles is known as **diffraction** of light. The intensity pattern produced is called a **diffraction pattern**.

The interpretation of the diffraction phenomena was given by Fresnel. According to him, the diffraction phenomenon is due to the interference of secondary wavelets originating from various points of the wavefront which are not blocked by the obstacle. At this stage it would be useful to understand the difference between the phenomena of interference and diffraction as follows:

- (i) In interference, interaction takes place between two wavefronts originating from two coherent sources, whereas in diffraction, the interaction is between secondary wavelets originating from different points of the same wavefront.

- (ii) In interference, all the maxima are of same intensity, but in diffraction pattern they are of varying intensity.
- (iii) The fringewidths in a diffraction pattern are never equal, whereas in the case of interference, the fringewidths could be equal at least in certain cases.
- (iv) In the interference pattern, the regions of minimum intensity are almost dark. It is not the case in diffraction.

4.2 TYPES OF DIFFRACTION

The diffraction phenomenon is divided into two categories—Fraunhofer diffraction and Fresnel diffraction. In Fraunhofer diffraction, the source of light and the screen are effectively at infinite distances from the obstacle or aperture which causes the diffraction, refer Fig. 4.1(a). As the source is at infinity, the waves reaching the obstacles are plane waves. As the screen is at infinity, interference pattern produced by secondary waves on the screen is due to parallel rays.

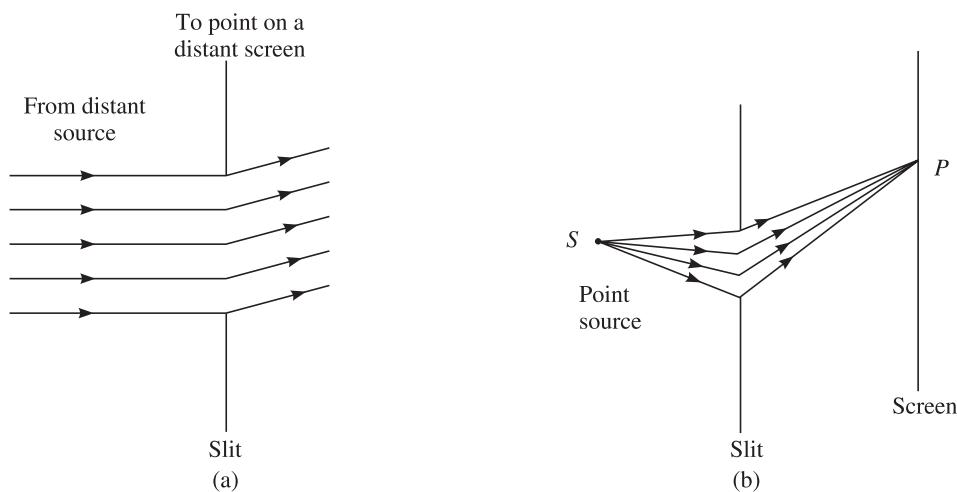


Fig. 4.1 (a) Fraunhofer diffraction (b) Fresnel diffraction.

In Fresnel diffraction, the source of light or the screen or both at finite distances from obstacle or aperture causing the diffraction. If the source is at finite distance, the incident wavefront is either spherical or cylindrical. In general, the calculation of intensity of the patterns on the screen is complex. The general experimental arrangement is shown in Fig. 4.1(b). Fresnel diffraction is more general. It includes Fraunhofer diffraction as a special case.

In practice, the arrangement for Fraunhofer diffraction is achieved by using two convergent lenses as shown in Fig. 4.2. The first lens makes the light beam parallel. The second lens effectively makes the screen receive the parallel beam of light.

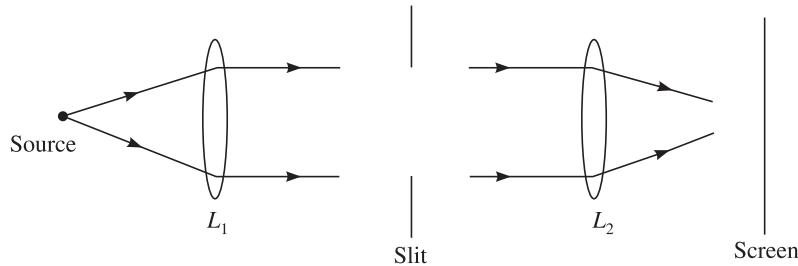


Fig. 4.2 Fraunhofer diffraction conditions produced by lenses.

4.3 FRAUNHOFER DIFFRACTION AT A SINGLE SLIT

Fraunhofer diffraction refers to diffraction effects produced by parallel collimated light. This condition can easily be realized in practice by using collimating lens, keeping the slit source at a distance equal to the focal length of the lens.

4.3.1 Experimental Arrangement

A slit is a rectangular aperture whose length is large compared to its breadth. Figure 4.3(a) gives the schematic representation of Fraunhofer diffraction by a slit of width \$AB = d\$, which is illuminated by a point source \$S\$. Lens \$L_1\$ render the rays from \$S\$ parallel. According to geometrical optics, lens \$L_2\$ should converge the parallel rays and produce an image at \$S'\$ in the focal plane of the lens. However, due to diffraction effects, the intensity outside \$S'\$ is not zero. The observed intensity distribution is shown in Fig. 4.3(b).

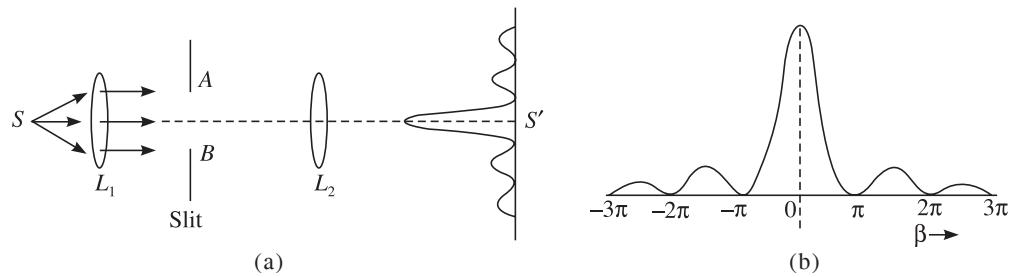


Fig. 4.3 Fraunhofer diffraction at a single slit (a) Experimental arrangement (b) Intensity pattern of single slit pattern.

4.3.2 Positions of Maximum and Minimum Intensities

To obtain the positions of intensity maxima and minima, one has to do a complete mathematical analysis. As the plane wave is incident on the slit \$AB\$, each point on the wave front is a source of secondary disturbance. The secondary waves which are at equal distances from the axis have to travel equal distance to reach \$S'\$. Hence, the path difference will be zero leading to the position of maximum intensity. For a semi-quantitative analysis, a parameter \$\beta\$ is defined by Eq. (4.1):

$$\beta = \frac{\pi d \sin \theta}{\lambda} \quad (4.1)$$

where θ is the angle which a ray that reaches outside s' makes with the ray through the centre of the slit that reaches s' . Detailed calculations Appendix II lead to the following expression for intensity distribution:

$$I = I_0 \frac{\sin^2 \beta}{\beta^2} \quad (4.2)$$

where I_0 represents the intensity at $\theta = 0$. From Eq. (4.2), it is evident that the intensity is zero when

$$\beta = n\pi, \quad n = \pm 1, \pm 2, \pm 3, \dots$$

$$I = I_0 \frac{\sin^2 n\pi}{n^2 \pi^2} = 0 \quad (4.3)$$

Substituting this value of β in Eq.(4.1), we get

$$d \sin \theta = n\lambda, \quad n = \pm 1, \pm 2, \dots \quad (4.4)$$

which is the condition for an intensity minimum. When $\beta = 0$, from Eq. (4.2), we have

$$I = I_0 \frac{\sin \beta}{\beta} \frac{\sin \beta}{\beta} = I_0$$

That is, the condition $\beta = 0$ corresponds to the central maximum. The separation between the first minimum on both sides can be taken as the width of the central bright band.

To determine the positions of other maxima, let us differentiate Eq. (4.2) with respect to β and set it to zero, we will obtain

$$\begin{aligned} \frac{dI}{d\beta} &= I_0 \left(\frac{2 \sin \beta \cos \beta}{\beta^2} - \frac{2 \sin^2 \beta}{\beta^3} \right) = 0 \\ \sin 2\beta - \frac{2 \sin \beta \cos \beta \tan \beta}{\beta} &= 0 \\ \sin 2\beta \left(1 - \frac{\tan \beta}{\beta} \right) &= 0 \end{aligned}$$

As $\beta \neq 0$ for noncentral maximum, $\sin 2\beta \neq 0$. Hence

$$(\beta - \tan \beta) = 0 \quad (4.5)$$

Hence, to get the condition for other maxima, one has to find the roots of the transcendental equation

$$\tan \beta = \beta \quad (4.6)$$

The solution of Eq. (4.6) can be obtained graphically. The graphical solution gives

$$\beta = 1.43\pi, 2.46\pi, \dots \quad (4.7)$$

These are referred to as the **first maximum**, **second maximum** and so on. The zero intensity positions are $\beta = \pm 1\pi, \pm 2\pi, \pm 3\pi, \dots$. The half way positions of zero intensity are $\beta = \pm 1.5\pi$,

$\pm 2.5\pi, \pm 3.5\pi, \dots$. It may be noted that the actual maxima are in between the zero intensity points, but slightly shifted towards the central band. For the first maximum

$$\frac{I}{I_0} = \left(\frac{\sin 1.43\pi}{1.43\pi} \right)^2 \cong 0.0496 \quad (4.8)$$

That is, the intensity of the first maximum is about 4.96% of the central maximum. Similarly the intensity of the second maximum is about 1.68% of the central maximum.

4.4 DIFFRACTION AT A CIRCULAR APERTURE

Diffraction of a plane wave by a circular aperture is of importance since most of the lenses we use are round. Airy was the first to work out the problem in 1835. Figure 4.4 shows the experimental arrangement for observing its Fraunhofer diffraction pattern. The plane wave incident on the circular aperture falls on the lens L_2 and the pattern is formed at the focal plane of the lens. The diffraction pattern consists of a series of maxima and minima in the form of concentric rings.

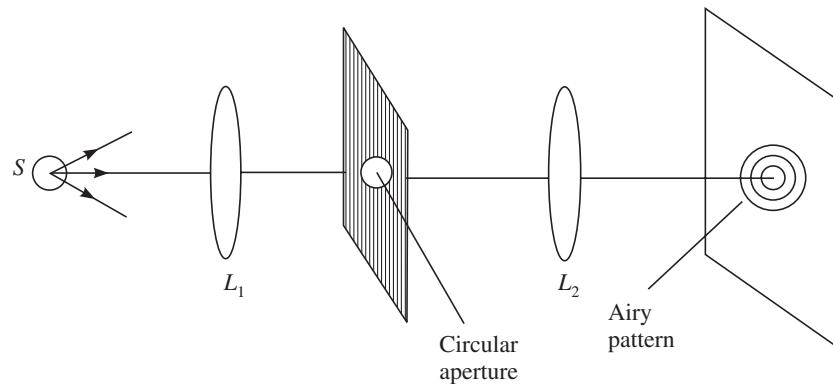


Fig. 4.4 Arrangement for observing Fraunhofer diffraction by a circular aperture.

The mathematical analysis of the problem is somewhat complicated. As in the case of diffraction by slit, the aperture is divided into a series of narrow strips of unequal length but of equal width. The resultant amplitude is calculated by taking the contributions from all these strips. The intensity distribution is similar to that obtained in the case of diffraction by a slit, but the dimensions are different. The dark rings appear when

$$\sin \theta = \frac{3.832\lambda}{\pi D}, \frac{5.136\lambda}{\pi D}, \frac{7.016\lambda}{\pi D}, \dots \quad (4.9)$$

where D is the diameter of the aperture, λ the wavelength of light used and θ is the angle of diffraction. The disk bounded by the first dark ring is often called **Airy disc**. For the first dark ring

$$\sin \theta = \frac{3.832\lambda}{\pi D} \equiv \frac{1.22\lambda}{D}, \dots \quad (4.9a)$$

The detailed mathematical analysis shows that about 84% of the energy is contained within the first dark ring. Thus, the airy disc is surrounded by alternate dark and bright rings, called the **airy pattern**. The central spot is very bright and the intensity of the rings decreases as one moves away from the centre (Fig. 4.5).



Fig. 4.5 Diffraction pattern due to diffraction at a circular aperture.

4.5 DIFFRACTION GRATING

4.5.1 Transmission and Reflection Gratings

Though David Rittenhouse invented diffraction grating in 1785, the important contributions came only later from Fraunhofer. Gratings are mainly of two types, the transmission grating and reflection grating. A system of equally spaced slits acts as a transmission grating. It is made by ruling a series of closely spaced parallel lines with a diamond point on an optically transparent sheet of material. The ruling on a reflecting surface such as aluminum can make the reflection grating. The transmission and reflection gratings are illustrated in Fig. 4.6. The distance between corresponding points of adjacent slits is called the **grating element** d . If the width of transparent portion is a and opaque portion is b , then $d = a + b$. If there are N number of lines (slits) per unit length, the grating element, is given by

$$d = \frac{1}{N} = \frac{1}{(a+b)} \quad (4.10)$$

Usually, the number N varies from 5000 to 7000 lines per centimetre.

Large-scale production of gratings are made by taking the cast of an actual grating on a transparent film like cellulose acetate. Cellulose acetate solution of proper strength is poured on the ruled surface and is allowed to dry. A strong thin film is formed which is detached from the parent grating and preserved by mounting it between two glass sheets. Holographic gratings, where one records the interference pattern of two plane or spherical waves, are also available.

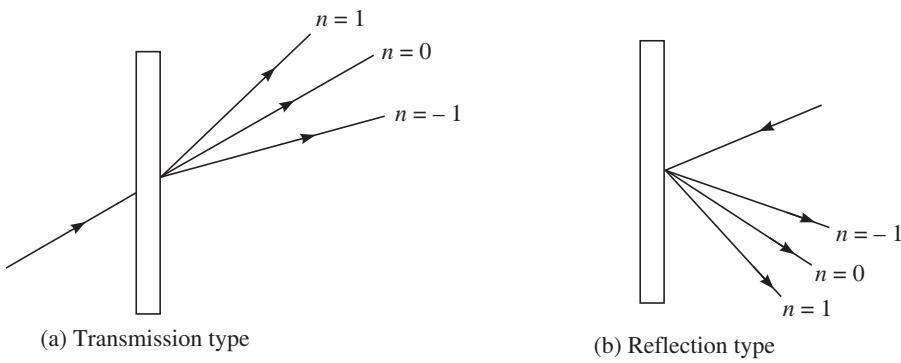


Fig. 4.6 Grating.

The number of lines per centimetre of holographic gratings are much larger compared to the ruled gratings.

4.5.2 Experimental Arrangement

The diffraction pattern exhibited by a grating is known as the **grating spectrum**. Figure 4.7 shows the experimental set up for the diffraction of a plane wave incident normally on a transmission grating. The rays are diffracted by the slits of the grating and the diffracted rays produce interference pattern in the screen. The secondary waves travelling in the same direction as that of the incident light reinforce at the point P [refer Fig. 4.8(a)]. This point will be the central bright maximum of the pattern. Secondary waves travelling in a direction inclined at an angle to the direction of incident light come to focus at points on both sides of P . Such points will be dark or bright depending on the path difference of secondary waves originating from two neighbouring slits. Usually, the grating is mounted on the prism table of the spectrometer, and therefore, the angle through which the rays are diffracted can be measured very easily. If we use white light as source, then the central maximum ($n = 0$) is also white. For higher orders, the angles of diffraction are different for different wavelengths, and therefore, various spectral components appear at different positions giving rise to a spectrum. By measuring the angles of diffraction for the different wavelengths one can determine the values of wavelengths.

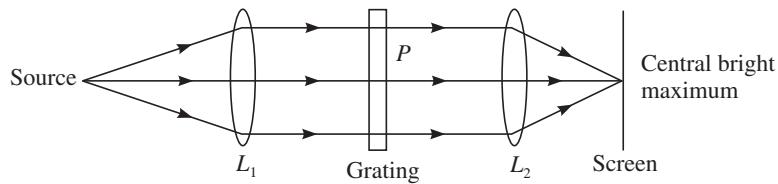


Fig. 4.7 Grating spectrograph.

A grating spectrograph in Fig. 4.7 shows secondary waves giving rise to the central bright maximum.

4.5.3 Grating Equation

Let AB , CD , and so on, be the ruled portion of the grating, and θ be the angle of diffraction corresponding to first order maximum [Fig. 4.8(b)]. From Fig. 4.8, we get

$$\angle CAF = \theta; \quad \sin \theta = \frac{CF}{AC} \quad (4.11)$$

Path difference between rays emerging from the corresponding points of adjacent slits is CF .

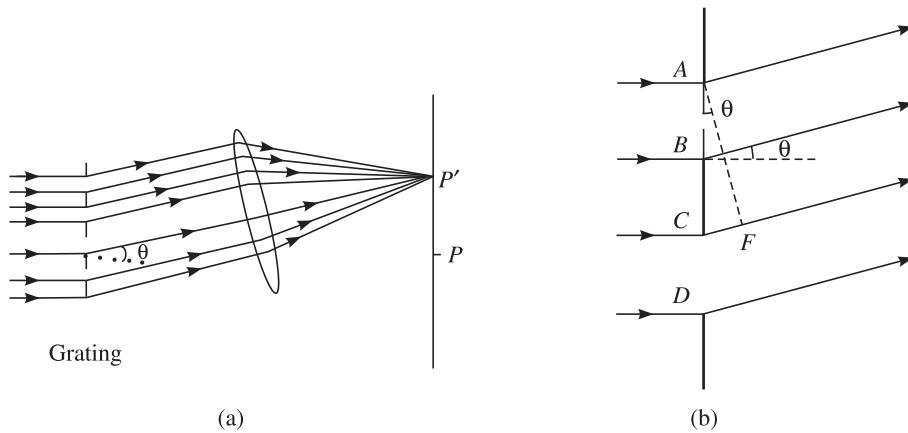


Fig. 4.8 Diffraction by a grating (a) Secondary waves travelling in a direction inclined at an angle θ . (b) Grating showing the diffracted rays and angle of diffraction θ .

Since $AC = d = \frac{1}{N}$, from Eq. (4.11), we get

$$CF = AC \sin \theta = d \sin \theta = \frac{\sin \theta}{N} \quad (4.12)$$

For the n th order principal maximum, the path difference must be equal to $n\lambda$. Hence,

$$\frac{\sin \theta}{N} = n\lambda, \quad n = 1, 2, 3, \dots$$

$$\sin \theta = nN\lambda, \quad n = 1, 2, 3, \dots \quad (4.13)$$

Equation (4.13) is called the **grating equation**. The angle θ cannot exceed 90° . Hence, the number of orders possible is limited by the value of N . A coarse grating, with only a few slits per unit length will give many orders. When white light is used, in a given order, the blue colour appears nearer to the axis and the red colour farther away from it since $\lambda_{\text{red}} > \lambda_{\text{blue}}$. The diffraction angle for the different wavelengths corresponding to the different orders are determined and the wavelengths calculated using Eq. (4.13).

4.5.4 Dispersive Power

As we have learnt for polychromatic sources each order of the spectrum (except the zeroth order) will consist of different wavelengths depending on the source. If $d\theta$ is the separation for wavelengths differing by $d\lambda$, the dispersive power ($d\theta/d\lambda$) can be obtained from the grating equation, Eq. (4.13). Differentiating Eq. (4.13) with respect to λ , we get

$$\cos \theta \frac{d\theta}{d\lambda} = nN \quad \text{or} \quad \frac{d\theta}{d\lambda} = \frac{nN}{\cos \theta} \quad (4.14)$$

Therefore, the dispersive power is:

- (i) Directly proportional to the number of lines per unit length of the grating
- (ii) Inversely proportional to the cosine of diffraction angle
- (iii) Directly proportional to the order of the spectrum

From the first one, it is evident that smaller the grating element, the larger will be the angular dispersion. From statement (iii) we have the result that the dispersive power in the second order is twice than that of the first order.

4.6 RESOLVING POWER

Often, it is important to know how close two wavelengths may be and still be distinguished as separate ones. The ability of an instrument to discriminate wavelengths λ and $\lambda + \Delta\lambda$ is

expressed by the ratio $\left(\frac{\lambda}{\Delta\lambda} \right)$ which is called the **resolving power** R of the instrument:

$$R = \frac{\lambda}{\Delta\lambda} \quad (4.15)$$

Figure 4.9 shows two spectral lines resolved to different extents. The dashed curves represent the individual lines the solid curves the resultant. In Fig. 4.9(a) the resultant intensity has only one peak, and hence, the two lines are unresolved. In Fig. 4.9(b) it is just resolved whereas in Fig. 4.9(c) the lines are well resolved.

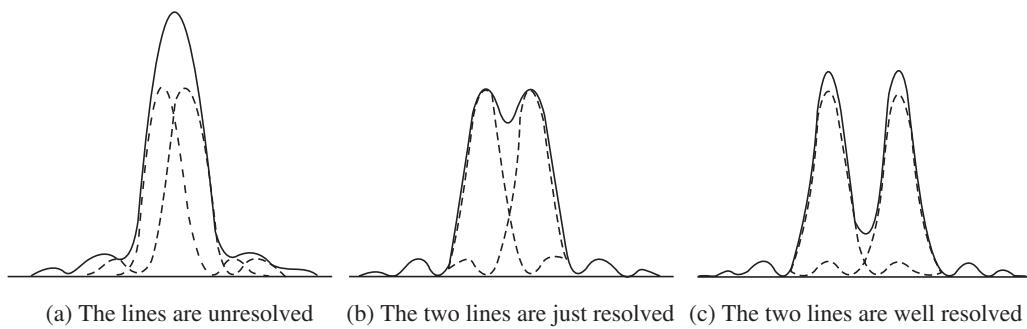


Fig. 4.9 Two spectral lines resolved to different extents.

4.6.1 Rayleigh's Criterion

Diffraction, limits the resolving power of an optical system. Different criteria are available to specify whether two spectral lines are just resolved or not. Of these, the Rayleigh criterion is the commonly used. If the central diffraction maximum of one source (say wavelength λ) coincides with the first minimum of the other source (wavelength $\lambda + \Delta\lambda$), then the two wavelengths λ and $\lambda + \Delta\lambda$ are said to be just resolved. Figure 4.9(c) illustrates this criterion of Rayleigh. Rayleigh criterion stated in this form may require slight modification while applying to different situations. It can be shown that the diffraction effect limits the minimum angle of resolution provided by a lens of diameter D at a wavelength λ to:

$$\theta_{\min} = \frac{1.22\lambda}{D} \quad (4.16)$$

This is illustrated by the images formed for two distant point objects by the objective of a telescope in Fig. 4.10.

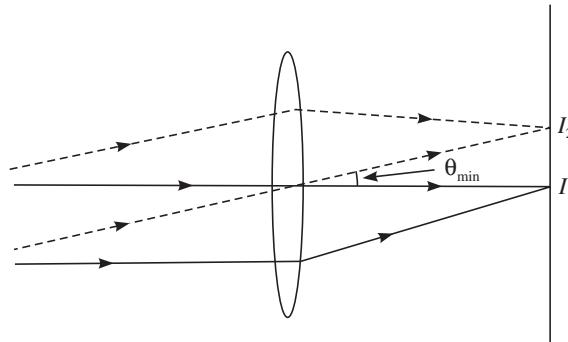


Fig. 4.10 The images of two distant point objects formed by the objective of a telescope.

4.7 RESOLVING POWER OF A GRATING

The resolving power of an instrument is defined by the equation

$$R = \frac{\lambda}{\Delta\lambda}$$

where $\Delta\lambda$ is the separation of two wavelengths which the grating can just resolve.

According to Rayleigh's criterion, the wavelengths λ and $\lambda + \Delta\lambda$ are said to be resolved if the principal maximum corresponding to wavelength $\lambda + \Delta\lambda$ falls on the first minimum of the wavelength λ . Let us denote the common diffraction angle by θ . It can be proved that between two principal maxima, we have minima with angles of diffraction corresponding to:

$$d \sin \theta = \frac{\lambda}{N_0}, \frac{2\lambda}{N_0}, \frac{3\lambda}{N_0}, \dots$$

where N_0 is the total number of lines in the grating exposed to the beam. If we are looking at the n th order spectrum, the first minimum after the n th order maximum for wavelength λ is given by

$$d \sin \theta = n\lambda + \frac{\lambda}{N_0} \quad (4.17)$$

The n th order principal maximum for the wavelength $(\lambda + \Delta\lambda)$ is given by

$$d \sin \theta = n(\lambda + \Delta\lambda) \quad (4.18)$$

From Eqs. (4.17) and (4.18), we have

$$\begin{aligned} n\lambda + \frac{\lambda}{N_0} &= n(\lambda + \Delta\lambda) \\ R &= \frac{\lambda}{\Delta\lambda} = nN_0 \end{aligned} \quad (4.19)$$

The following results are obtained:

- (i) The resolving power of a grating is directly proportional to the total number of lines in the grating exposed to the beam.
- (ii) As in the case of dispersive power, the resolving power is proportional to the order of the spectrum.

4.8 GRATING AND PRISM SPECTRA—COMPARISON

In higher order diffraction grating with a composite light source, different wavelengths appear at different positions giving rise to a spectrum. In the same way, different spectral lines of a composite source can be dispersed into a spectrum with a prism. The salient features of grating and prism spectra are as follows:

- (i) Grating spectrum arises due to diffraction, whereas the prism spectrum is due to the dispersion of light.
- (ii) In the prism spectrum, the deviation for lower wavelengths (violet end) are more than that for higher wavelengths (red end). However, reverse is the case in grating spectra since $2d \sin \theta = n\lambda$.
- (iii) The grating spectrum consists of number of orders, whereas the prism spectrum is a single set.
- (iv) Grating spectrum is weak as compared to the prism spectrum since part of the incident light is obstructed by the opaque positions of the grating, and the remaining incident intensity is distributed among various orders.
- (v) The spectral lines in a grating spectrum will be better resolved than in a prism spectrum, since the resolving power of a grating is very high as compared to that of a prism.
- (vi) The dispersive power of a grating is fairly constant for a given order. Hence, the spectral lines are evenly distributed in grating spectrum. In the case of a prism, it is not uniform, and is greater towards high frequencies (violet end).
- (vii) The mathematical relation connecting wavelength and deviation is a simple in the case of a grating (grating equation). However, it is much more complex in the case of a prism.

In general, grating spectrum gives more accurate information about the system under consideration than the information given by the prism spectrum. Hence, in sophisticated spectrophotometers grating is used as the dispersing element.

4.9 RESOLVING POWER OF A TELESCOPE

The resolving power of a telescope is its ability to form separate images for two distant point objects close to each other. It is measured in terms of the angle subtended at the objective of the telescope by the two nearby distant objects whose images are just resolved by the telescope. This angle is the limit of resolution of the telescope. Smaller the value of the angle, higher is the resolving power of the telescope. Hence, the reciprocal of the limit of resolution is taken as the resolving power of the telescope.

Let A and B be two nearby distant objects sending light of wavelength λ , (refer Fig. 4.11) As the sources are far away from the objective, the wave front falling on it will be plane. The objective of the telescope acts as a circular aperture, and hence, the image of a distant object will be a diffraction pattern consisting of a central bright disc surrounded by alternate dark and bright rings. The central maximum of the diffraction pattern of the distant object A falls at P and the first minimum at Q in the focal plane of the lens. For the first minimum, from Eq. (4.9a), we have

$$\sin \theta = \frac{1.22\lambda}{D} \quad (4.20)$$

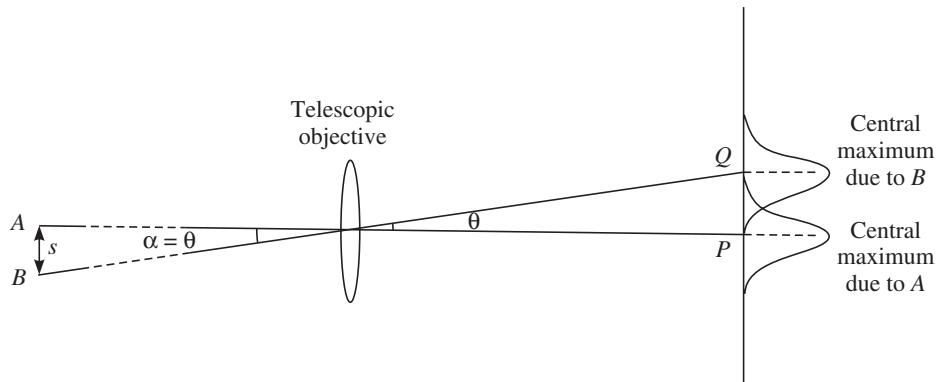


Fig. 4.11 The central maxima of two distant objects A and B produced by a telescope objective.

where D is the diameter of the objective lens. In general $D \gg \lambda$, and hence, $\sin \theta \approx \theta$. Hence,

$$\theta = \frac{1.22\lambda}{D} \quad (4.21)$$

If the central maximum of the nearby distant object B falls at Q , the diffraction pattern will be just resolved (Rayleigh's condition). If A and B subtend the angle α at the objective then for the just resolution of the images of A and B , we must have

$$\alpha = \theta_{\min} = \frac{1.22\lambda}{D} \quad (4.22)$$

The limit of resolution, is thus, proportional to the wavelength of the light used and inversely proportional to the diameter of the telescope objective. Hence, to minimise θ_{\min} lens of large, diameter has to be used. Resolving power (*R.P.*) of the telescope is then given by

$$R.P = \frac{1}{\theta_{\min}} = \frac{D}{1.22\lambda} \quad (4.23)$$

4.10 RESOLVING POWER OF A MICROSCOPE

As per the Rayleigh criterion, two nearby point objects are just resolvable if they are separated by an angle θ (see Fig. 4.9). the condition is given by Eq. (4.16) as

$$\theta = \frac{1.22\lambda}{D} \quad (4.24)$$

This relation is valid for microscope also. For a microscope it is convenient to specify the distance between the two objects that are just resolvable by s (see Fig. 4.11).

Since objects are normally placed near the focal point of the microscope objective, we obtain

$$\theta = \frac{s}{f} \quad \text{or} \quad s = \theta f \quad (4.25)$$

where f is the focal length of the microscope objective. The distance s is the minimum separation of two nearby objects that can be just resolved. Hence, for the limit of resolution, we obtain

$$s = \frac{1.22\lambda f}{D} \quad (4.26)$$

This equation is often written in terms of the angle of acceptance α of the objective lens which is defined in Fig. 4.12. It can be proved that in terms of α

$$s \equiv \frac{1.22\lambda}{2\sin\alpha} = \frac{0.61\lambda}{\sin\alpha} \quad (4.27)$$

Hence, the resolving power *R.P* is given by

$$R.P = \frac{\sin\alpha}{0.61\lambda} \quad (4.28)$$

The resolving power can be increased by placing a drop of oil that encloses the object and the front surface of the objective. This is called an **oil immersion objective**. In that case

$$R.P = \frac{n \sin\alpha}{0.61\lambda} \quad (4.29)$$

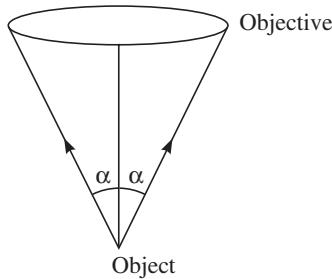


Fig. 4.12 Objective lens of a microscope showing the angle of acceptance α .

The oil has a typical value of the refractive index $n \approx 1.5$. Thus, oil immersion increases the resolving power by 50% or more.

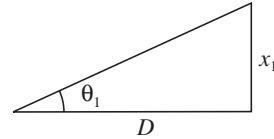
SOLVED EXAMPLES

Example 4.1 Parallel light of wavelength 6563\AA is incident normally on a slit 0.385 mm wide. A lens with a focal length of 50.0 cm located just behind the slit brings the diffraction pattern to focus on a white screen. Find the distance from the centre of the principal maximum to (a) the first minimum and (b) the fifth minimum.

Solution: The positions of a minima are given by $d \sin \theta = n\lambda$, where d is the slit width and n is the order of the minimum.

$$(a) \sin \theta_1 = \frac{n\lambda}{d}$$

$$= \frac{6563 \times 10^{-8} \text{ cm}}{0.0385 \text{ cm}} = 0.0017$$



Since θ_1 is very small,

$$\sin \theta_1 \approx \frac{x_1}{D} \quad \text{or} \quad x_1 = \sin \theta_1 \times D$$

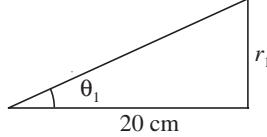
$$x_1 = 0.0017 \times 50 \text{ cm} = 0.085 \text{ cm} = 0.85 \text{ mm}$$

$$(b) \sin \theta_2 = 5 \times 0.0017 = 0.0085$$

$$x_2 = 0.0085 \times 50 \text{ cm} = 0.425 \text{ cm} = 4.25 \text{ mm}$$

Example 4.2 Calculate the radii of the first two dark rings of the diffraction pattern produced by a circular aperture of diameter 0.04 cm at the focal plane of a convex lens of focal length 20 cm . Wavelength of light used is 6000\AA .

Solution: For the first dark ring,

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


$$\sin \theta \approx \theta = \frac{1.22 (6000 \times 10^{-8} \text{ cm})}{0.04 \text{ cm}} = 1.83 \times 10^{-3}$$

$$\theta = \frac{r_1}{20 \text{ cm}}$$

$$r_1 = 1.83 \times 10^{-3} \times 20 \text{ cm} = 3.66 \times 10^{-2} \text{ cm}$$

For the second dark rings,

$$\sin \theta = \theta = \frac{5.136 \times \lambda}{\pi D} = \frac{5.136}{\pi} \frac{(6000 \times 10^{-8} \text{ cm})}{0.04 \text{ cm}} = 2.45 \times 10^{-3}$$

$$r_2 = 2.45 \times 10^{-3} \times 20 \text{ cm} = 4.9 \times 10^{-2} \text{ cm}$$

Example 4.3 At what angle will 650 nm light produce a second order maximum when falling on a grating whose slits are 1.2×10^{-3} cm apart?

Solution: The grating equation is

$$\sin \theta = nN\lambda = \frac{n\lambda}{d}$$

$$\sin \theta = \frac{2(650 \times 10^{-9} \text{ m})}{1.2 \times 10^{-5} \text{ m}} = 0.1083$$

$$\theta = 6.22^\circ$$

Example 4.4 What is the highest spectral order that can be seen if a grating with 6000 lines/cm is illuminated by a source of 650 nm wavelength? (Assume normal incidence.)

Solution: Since the maximum possible value of $\sin \theta = 1$, from Eq. (4.13), we get

$$n = \frac{\sin \theta}{N\lambda} = \frac{1}{(6 \times 10^5 \text{ lines/m})(650 \times 10^{-9} \text{ m})} = 2.56$$

Hence, the highest spectral order that can be seen is 2.

Example 4.5 White light containing wavelengths from 400 nm to 750 nm strikes a grating having 4000 lines/cm. Show that the blue line at $\lambda = 450$ nm of the third order overlaps the red at 700 nm of the second order.

Solution: The blue of the third order occurs at an angle θ which is given by

$$\sin \theta = nN\lambda = 3(4 \times 10^5 \text{ lines/m})(450 \times 10^{-9} \text{ m}) = 0.540$$

Red of the second order occurs at

$$\sin \theta = nN\lambda = 2(4 \times 10^5 \text{ lines/m}) (700 \times 10^{-9} \text{ m}) = 0.560$$

Since $\sin \theta$ for the two are almost equal, the two orders overlap.

Example 4.6 White light containing wavelengths from 400 nm to 750 nm falls on a grating with 6000 lines/cm. How wide is the first order spectrum on a screen 2 m away?

Solution: For 400 nm,

$$\sin \theta = \left(6 \times 10^5 \frac{\text{lines}}{\text{m}} \right) (400 \times 10^{-9} \text{ m}) = 0.24$$

$$\theta = 13.88^\circ$$

For 750 nm,

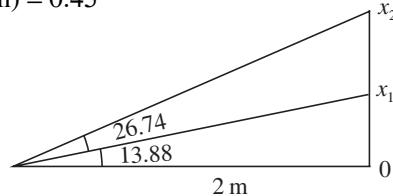
$$\sin \theta = \left(6 \times 10^5 \frac{\text{lines}}{\text{m}} \right) (750 \times 10^{-9} \text{ m}) = 0.45$$

$$\theta = 26.74^\circ$$

$$x_1 = 2 \tan 13.88^\circ = 0.4942 \text{ m}$$

$$x_2 = 2 \tan 26.74^\circ = 1.0076 \text{ m}$$

$$x_2 - x_1 = 0.513 \text{ m} = 51.3 \text{ cm}$$



Example 4.7 What is the least separation between wavelengths that can be resolved near 640 nm in the second order, using a diffraction grating that is 5 cm wide and ruled with 32 lines per millimetre?

Solution: Number of lines in the grating = 5×320

From Eq. (4.19), we get

$$\Delta\lambda = \frac{\lambda}{nN_0} = \frac{640 \text{ nm}}{2 \times 5 \times 320} = 0.2 \text{ nm}$$

Example 4.8 A circular lens, with diameter 3.2 cm and focal length 24 cm forms images of distant point objects in the focal plane of the lens. If the wavelength of light used is 550 nm, what angular separation must two distant point objects have to satisfy Rayleigh's criterion. What is the separation Δx of the centres of the images in the focal plane?

Solution: From Eq. (4.16), we get

$$\theta_{\min} = \frac{1.22\lambda}{D} = \frac{1.22(550 \times 10^{-9} \text{ m})}{3.2 \times 10^{-2} \text{ m}} = 2.1 \times 10^{-5} \text{ rad}$$

$$\frac{\Delta x}{f} = \theta_{\min} \quad \text{or} \quad \Delta x = \theta_{\min} f$$

$$\Delta x = (2.1 \times 10^{-5} \text{ rad}) (0.24 \text{ m}) = 5 \times 10^{-6} \text{ m} = 5 \mu\text{m}$$

Example 4.9 A telescope objective having diameter 20 cm is directed towards two points at a distance of 6 km from the objective lens. If the wavelength of light used is 550 nm, evaluate the distance between the two points.

Solution: From Eq. (4.21), we obtain

$$\begin{aligned}\theta &= \frac{1.22\lambda}{D} = \frac{1.22(550 \times 10^{-9} \text{ m})}{20 \times 10^{-2} \text{ m}} \\ &= 3.355 \times 10^{-6}\end{aligned}$$

Distance between the two points is given by

$$\begin{aligned}x &= \theta \times 6 \text{ km} = (3.355 \times 10^{-6})(6 \times 10^3 \text{ m}) \\ &= 20.13 \times 10^{-3} \text{ m} = 20.13 \text{ mm}\end{aligned}$$

REVIEW QUESTIONS

- 4.1 Explain the phenomenon of diffraction of light.
- 4.2 Distinguish between Fraunhofer and Fresnel diffractions.
- 4.3 Explain the difference between interference and diffraction phenomena.
- 4.4 Explain how a diffraction grating produces the spectrum of a given light.
- 4.5 Explain Rayleigh's criterion for resolution.
- 4.6 What do you understand by a grating element? Write down the grating equation in terms of grating element.
- 4.7 Explain the effect of increasing the number of lines in a grating on the diffraction pattern.
- 4.8 Define the dispersive power of a grating. State the expression for it.
- 4.9 What do you understand by the resolving power of a grating? State the expression for it.
- 4.10 Compare grating and prism spectra.
- 4.11 Using grating equation, obtain an expression for the maximum number of orders that can be observed with a plane transmission grating.
- 4.12 Define the resolving power of a telescope. How does it depend on the wavelength of light used?
- 4.13 Explain the dependence of the resolving power of a microscope on the acceptance angle of its objective lens.
- 4.14 What is Fraunhofer diffraction? Discuss Fraunhofer diffraction at a single slit.
- 4.15 Distinguish between Fraunhofer and Fresnel diffractions. Discuss the Fraunhofer diffraction due to a circular aperture.

- 4.16** Give the theory of plane transmission grating for normal incidence. How would you use it to determine the wavelength of the prominent lines in the mercury spectrum?
- 4.17** Describe the method of determining the wavelength of sodium light using a transmission grating. Derive the necessary formula.
- 4.18** Define resolving power of a grating. Derive an expression for it.

PROBLEMS

- 1.** A grating produces a third order band at an angle of 22° . If the number of lines in the grating is 2500 per centimetre, what is the wavelength of the light used?
- 2.** A 10000 lines/cm grating gives three lines in the first order spectrum at angles of 31.2° , 36.4° and 47.5° . Calculate the wavelength of these lines.
- 3.** How many lines per centimetre does a grating have, if the third order occurs at 25° angle for 500 nm light?
- 4.** The first order line of 650 nm light falling on a diffraction grating is observed at 12° angle. How far apart are the slits? At what angle will the 4th order be observed?
- 5.** What is the highest order spectrum which may be seen with light of wavelength 6000 Å by means of a diffraction grating having 5000 lines/cm?
- 6.** Only two spectral orders are visible on either side of the central maximum when 400 nm light is passed through a diffraction grating. What is the maximum number of lines per centimetre for the grating?
- 7.** How many lines must a grating have, used in the second order and near 550 nm, to resolve two lines 0.1 Å apart?
- 8.** Plane waves of $\lambda = 4340 \text{ \AA}$ fall on a slit, then passes through a lens of focal length of 85 cm. If the central band of the diffraction pattern has a width of 2.45 mm, find the width of the slit (Hint: Width of the central bright band = Separation between the first minima on both sides).
- 9.** In a Fraunhofer diffraction due to a single slit of width 0.6 mm, the screen is 1 m away from the slit. If the wavelength of light used is 6000 Å, calculate the width of the central band.
- 10.** Find the diameter of the airy disk in the focal plane of a refracting telescope having an objective with a focal length of 1.0 m and a diameter of 10 cm. Assume the wavelength as $5.5 \times 10^{-5} \text{ cm}$.
- 11.** Light from a sodium lamp is allowed to fall normally on a grating having a width of 1.5 cm and 8000 lines. Find the angular separation of the 589 and 589.6 nm lines of sodium in the first order spectrum.

C H A P T E R

5

Polarization

The wave nature of light was clearly established by the phenomena of interference and diffraction. These phenomena failed to explain whether the light waves are longitudinal or transverse or whether the vibrations are linear, circular, or elliptical. These important factors form the subject matter of polarization which we discuss in this chapter. The phenomenon of polarization establishes the transverse nature of light waves.

5.1 INTRODUCTION

When a wave passes through a medium such as water or a rope, the particles of the medium are set into vibrations (Section 1.1). When these vibrations are perpendicular to the direction of propagation, the wave is said to be a **transverse wave**. When the vibrations travel parallel to the direction of propagation, it is referred to as a **longitudinal wave**. In the case of transverse wave motion, the particles of the medium can vibrate in an infinite number of planes. Hence, two similar transverse waves may differ from one another because of their different directions of vibrations. In a transverse wave, if the vibrations are in all possible directions, it is an **unpolarized wave**. When the vibrations of the particles of the medium are confined to a single plane, the wave is said to be **polarized**.

In unpolarized light, the electric field **E** vibrates in all directions perpendicular to the direction of propagation. Figure 5.1(a) represents unpolarized light, and Fig. 5.1(b) represents polarized light travelling to the right with vibrations in the plane of the paper. A plane polarized light travelling to the right with vibrations normal to the plane of the paper is represented in Fig. 5.1(c).

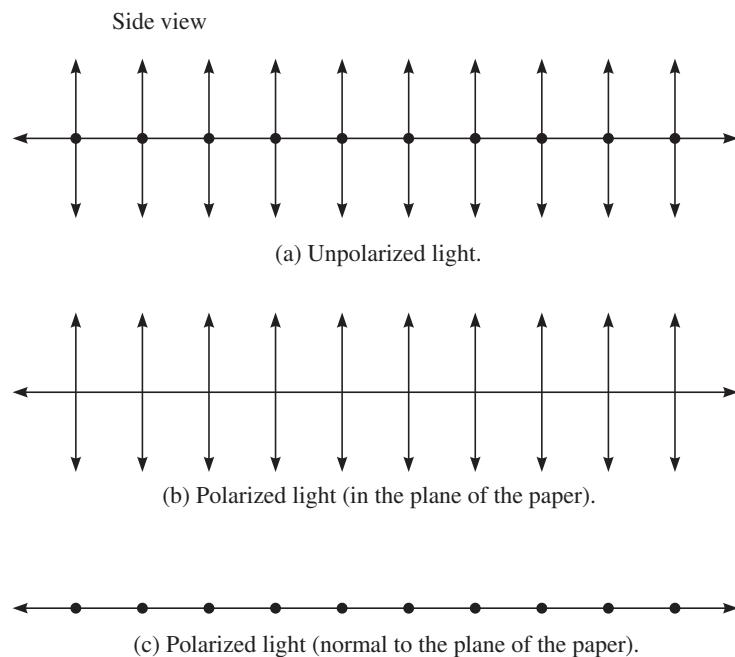


Fig. 5.1 Pictorial representation of unpolarized light, polarized light- E in the plane of the paper and polarized light- E normal to the plane of the paper.

5.2 NATURE OF LIGHT

Consider a tourmaline crystal A cut parallel to its crystallographic axis and held in the path of ordinary light such that light falls normally on it as shown in Fig. 5.2(a). If the crystal A is rotated about SO , there is no change in intensity of the transmitted beam. Now, another similar crystal B is placed in the path such that the axis of B is parallel to the axis of A . Then all the light transmitted by A will be completely transmitted by B .

Now, if crystal B is rotated about the beam axis, the intensity of light coming out of B decreases and becomes zero when the axes of A and B are perpendicular to each other. If B is rotated further, the intensity again increases and becomes maximum when both axes are parallel again. From the experiment, it is evident that the light coming out of the crystal A has acquired some property which prevents it from being transmitted by the second crystal when its axis is perpendicular to the first one. That is, the light coming out of crystal A has its oscillations confined to only one direction, the direction perpendicular to the direction of propagation of light. Hence, the light polarized by A will pass completely through B only if its axis is parallel to vibrations of light falling on B . It will be completely stopped if the axis of B is perpendicular to the vibrations of light falling on it [refer Fig. 5.2(b)]. This proves that light waves are transverse waves.

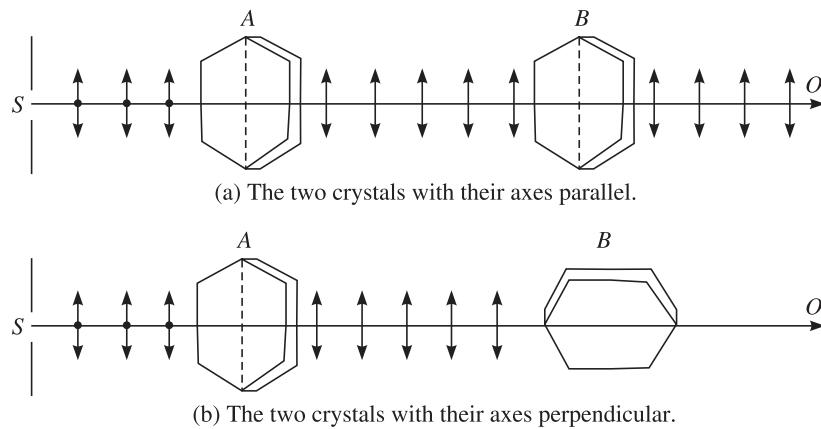


Fig. 5.2 Transverse wave nature of light. S – Unpolarized light source, A and B are tourmaline crystals.

5.3 TYPES OF POLARIZED LIGHT

Electromagnetic waves are radiation fields with the electric field vector **E** and the magnetic field vector **B** perpendicular to each other, and to the direction of propagation. As already pointed out, for studying optical effects we are primarily concerned with the electric field vector **E**.

5.3.1 Linearly Polarized Light

An electromagnetic wave (light wave) is said to be **linearly polarized** or **plane polarized** if the plane containing the electric field vector is constant. That is, the electric field vector at any point will be oscillating in a fixed direction perpendicular to the direction of propagation. Projection of the wave on a plane intercepting the axis of propagation gives a line, hence the term linearly polarized. The plane containing the electric vector **E** and the direction of propagation is called the **plane of vibration**. The plane containing the magnetic field vector **B**, which is perpendicular to the plane containing the electric vector, is the **plane of polarization**. Figure 5.3 represents a linearly polarized wave propagating along *x*-axis with the electric vector **E** (along *y*-axis) and the magnetic vector **B** (along *z*-axis) oscillating perpendicular to the direction of propagation.

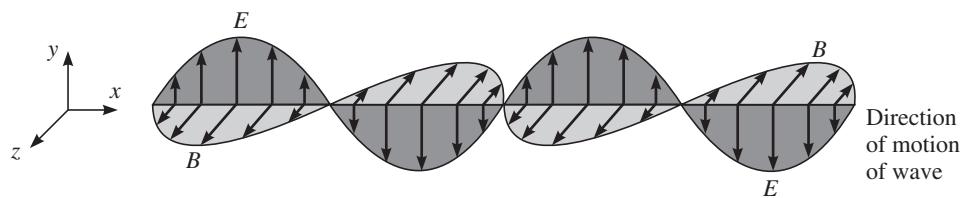


Fig. 5.3 A plane polarized electromagnetic wave that propagates along the *x*-axis. The wave is polarized in the *yz*-plane.

5.3.2 Circularly Polarized Light

In circular polarization, the electric field vector no longer oscillates in a plane as in linear polarization. However, the electric field vector is constant in magnitude and proceeds in the form of a helix around the axis of propagation [Fig. 5.4(a)]. Within one wavelength, the \mathbf{E} vector completes one revolution. When looking towards the light source, if the tip of the vector rotates clockwise as in Fig. 5.4(b) the light is said to be **right circularly polarized** and if it rotates anticlockwise it is **left circularly polarized**.

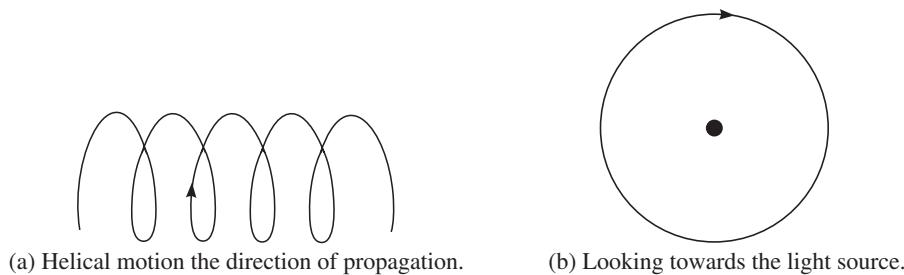


Fig. 5.4 Right circularly polarized light.

5.3.3 Elliptically Polarized Light

Elliptical polarization is the most general type of polarization, which stands between linear and circular polarizations. Linear and circular are the two extremes of elliptical polarization, as shown in Fig. 5.5. In this polarization, the tip of the vector proceeds in the form of a flattened helix. The vector rotates and its magnitude changes.

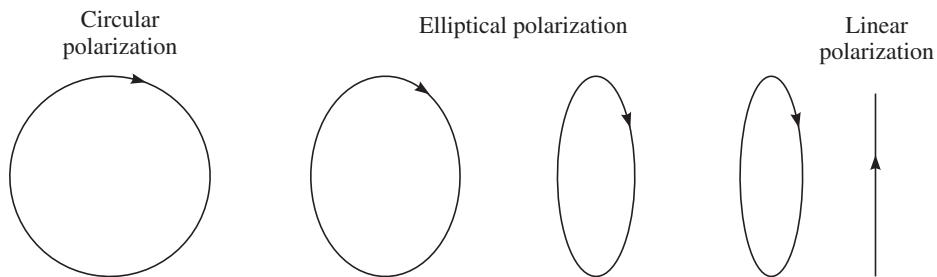


Fig. 5.5 Circular, elliptical (flattened to different levels) and linear polarization.

5.4 TYPES OF POLARIZED LIGHT—ANALYTICAL TREATMENT

In this section, we shall show that the polarized light is, in general, elliptical and the linearly and circularly polarized ones are its special cases. For a wave travelling in the z -direction, the electric field components in the x and y -directions are:

$$E_x = E_1 \sin \omega t \quad (5.1)$$

$$E_y = E_2 \sin(\omega t + \delta) \quad (5.2)$$

where E_1 and E_2 are the amplitudes of E_x and E_y , respectively. The angle δ is the time phase angle by which E_y leads E_x . Expanding the sine term in Eq. (5.2), we get

$$E_y = E_2 (\sin \omega t \cos \delta + \cos \omega t \sin \delta) \quad (5.3)$$

From Eq. (5.1), we have

$$\sin \omega t = \frac{E_x}{E_1} \quad \text{and} \quad \cos \omega t = \sqrt{1 - \left(\frac{E_x}{E_1}\right)^2} \quad (5.4)$$

Substituting the values of $\sin \omega t$ and $\cos \omega t$ in Eq. (5.3), we obtain

$$\frac{E_y}{E_2} = \frac{E_x}{E_1} \cos \delta + \sqrt{1 - \left(\frac{E_x}{E_1}\right)^2} \sin \delta \quad (5.5)$$

$$\frac{E_y}{E_2} - \frac{E_x}{E_1} \cos \delta = \sqrt{1 - \left(\frac{E_x}{E_1}\right)^2} \sin \delta \quad (5.6)$$

Squaring and rearranging, we get

$$\begin{aligned} \frac{E_y^2}{E_2^2} - \frac{2E_x E_y}{E_1 E_2} \cos \delta + \frac{E_x^2}{E_1^2} \cos^2 \delta &= \sin^2 \delta - \frac{E_x^2}{E_1^2} \sin^2 \delta \\ \frac{E_x^2}{E_1^2} + \frac{E_y^2}{E_2^2} - \frac{2E_x E_y}{E_1 E_2} \cos \delta &= \sin^2 \delta \end{aligned} \quad (5.7)$$

which is the general equation of an ellipse. We shall now consider some special cases.

(i) When $\delta = 0$; $\sin \delta = 0$, $\cos \delta = 1$. From Eq. (5.7), we obtain

$$\begin{aligned} \frac{E_x^2}{E_1^2} + \frac{E_y^2}{E_2^2} - \frac{2E_x E_y}{E_1 E_2} &= 0 \\ \left(\frac{E_x}{E_1} - \frac{E_y}{E_2}\right)^2 &= 0 \quad \text{or} \quad \left(\frac{E_x}{E_1} - \frac{E_y}{E_2}\right) = 0 \\ E_y &= \left(\frac{E_2}{E_1}\right) E_x \end{aligned} \quad (5.8)$$

This is the equation of a straight line and the wave is said to be **linearly polarized**.

(ii) When $\delta = \frac{\pi}{2}$ and $E_1 \neq E_2$; $\cos \delta = 0$, $\sin \delta = 1$. From Eq. (5.7), we obtain

$$\frac{E_x^2}{E_1^2} + \frac{E_y^2}{E_2^2} = 1 \quad (5.9)$$

which represents the equation of a symmetrical ellipse, and the wave is said to be **elliptically polarized**.

(iii) When $\delta = \frac{\pi}{2}$ and $E_1 = E_2 = E$. From Eq. (5.9), we have

$$E_x^2 + E_y^2 = E^2 \quad (5.10)$$

which is the equation of a circle. Here, the wave is said to be **left circularly polarized**. When $\delta = -\pi/2$, the wave is **right circularly polarized**.

5.5 POLARIZATION BY REFLECTION

In 1808, Malus discovered this simplest method of producing plane polarized light. When ordinary light is incident on the surface of any transparent material (water, glass, etc.), he found that the reflected beam is partially polarized. He also found that the degree of polarization depends on the angle of incidence. The reflected light is maximum (almost completely) polarized when the angle between the reflected and refracted rays is 90° (Fig. 5.6).

5.5.1 Brewster's Law

From Fig. 5.6, we get

$$i + 90^\circ + r = 180^\circ \quad \text{or} \quad r = 90^\circ - i$$

By Snell's law

$$\begin{aligned} \mu_1 \mu_2 &= \frac{\mu_2}{\mu_1} = \frac{\sin i}{\sin r} = \frac{\sin i}{\sin(90^\circ - i)} = \frac{\sin i}{\cos i} \\ \tan i &= \frac{\mu_2}{\mu_1} \end{aligned} \quad (5.11)$$

which is **Brewster's law**. The angle

$$i = \tan^{-1} \left(\frac{\mu_2}{\mu_1} \right) \quad (5.12)$$

is called **Brewster's angle of maximum polarization** on reflection or simply **polarizing angle**.

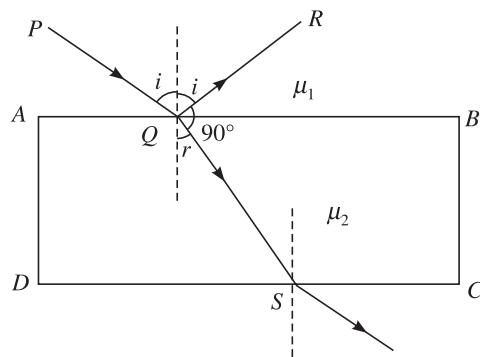


Fig. 5.6 Polarization by reflection.

The polarizing angle, as we know, depends on the reflecting substance and on the wavelength of light used. The value of polarizing angle is 57.5° for air–glass reflection, 53° for air–water reflection.

5.6 MALUS LAW

Consider a plane polarized light from a polarizer passing through an analyzer. Malus law states that the intensity (I) of the transmitted light through the analyzer varies as the square of the cosine of the angle between the transmission plane of the analyzer and the transmission plane of the polarizer. That is

$$I = I_0 \cos^2 \theta \quad (5.13)$$

where I_0 is the intensity of the incident polarized light on the analyzer. When $\theta = \pi/2$, $I = 0$. In that case the plane of the analyzer is perpendicular to the plane of the polarizer and no light is transmitted.

5.7 DOUBLE REFRACTION

Let us perform an experiment. Make a dot on a paper with a pen. Place a crystal of calcite over the dot and look through. Two images of the dot are seen. When the crystal is rotated about a vertical axis with the eye above the crystal, one image remains stationary while the other moves around it. The fixed image is called the **ordinary image** and the other one is called the **extraordinary image**.

In the same way when a ray of unpolarized light is incident normally on the surface of a calcite rhomb, it breaks up into two rays, one of which passes without deviation and the other deviates within the crystal, Fig. 5.7. Since the two opposite faces of a calcite crystal are always parallel to each other, the two refracted rays always emerge parallel to the incident light, and therefore, parallel to each other. The first one obeys the laws of refraction for all angles of incidence of the light and is called the **ordinary ray** or **o-ray**. The second one, which is refracted even at normal incidence, is known as the **extraordinary ray** or **e-ray** and for it Snell's law is not valid. The velocity of the ordinary ray is the same in all directions, whereas that of the e-ray is different in different directions. Along a particular direction of the crystal both the e-ray and o-ray have the same velocity. That is, there is no double refraction along that direction. That particular direction is called the **optic axis**.

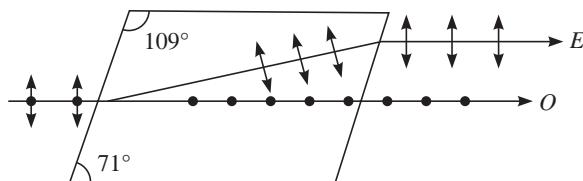


Fig. 5.7 Polarization by double refraction.

When the emergent rays are examined by an analyzer, it is found that they are linearly polarized, but is found that it is at right angle to each other. On one hand, the vibrations of

the ordinary ray will be perpendicular to the optic axis. On the other hand, vibrations of the extraordinary ray will be parallel to the optic axis. The effect of the calcite crystal, is therefore, to break up the light into two linearly polarized beams whose planes of vibrations are mutually perpendicular. This phenomenon of splitting of a light ray into two refracted rays is called **double refraction or birefringence**. A substance that exhibits different optical properties in different directions is called an **optically anisotropic** substance.

The refractive index of the crystal with respect to the ordinary ray is known as **ordinary refractive index** μ_o , which is a constant since the velocity of the o-ray is the same in all directions. For calcite, μ_o is 1.658 with respect to sodium light. The refractive index corresponding to the extraordinary ray μ_e is found to vary with the angle of incidence. It is maximum along the optic axis and is equal to μ_o . Its value is minimum (1.486) at right angles to the optic axis.

The o-ray has the same velocity in all directions, and therefore, if a point source of light is there within the crystal, the wavefronts would advance outward in the form of spheres. In the case of e-ray, the wavefronts would be ellipsoid of revolution since its velocity is different in different directions.

5.7.1 Uniaxial and Biaxial Crystals

Doubly refracting crystals are classified as uniaxial or biaxial. Uniaxial crystals have only one optic axis. Familiar examples of uniaxial crystal are calcite, quartz and tourmaline. In biaxial crystals there are two optic axes. These two axes make a certain angle with each other which is characteristic of the crystal. Mica and borax are biaxial crystals.

The uniaxial crystals are further divided into two—positive uniaxial and negative uniaxial. In negative uniaxial the velocity of e-ray is greater than that of the o-ray [see Fig. 5.8(a)], whereas the reverse is the case in positive uniaxial crystal [Fig. 5.8(b)]. Some of the general features of the two types are as follows:

Negative crystal

- (i) The velocity of the e-ray is different in different directions with the minimum value along optic axis and the maximum value in a direction perpendicular to the optic axis. The velocity of the o-ray is the same in all directions and is less than that for the e-ray.
- (ii) The refractive index for the e-ray varies, with the maximum value along the optic axis and the minimum value in a direction perpendicular to it. That is, $\mu_e < \mu_o$ with the condition that the maximum value of $\mu_e = \mu_o$.

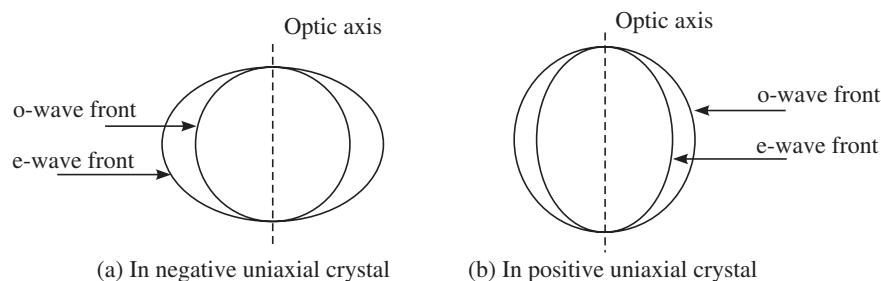


Fig. 5.8 Wavefronts.

- (iii) The wavefront for the o-ray lies inside the wavefront for the e-ray.

Examples: Calcite, tourmaline, and dolomite.

Positive crystals

- (i) Velocity of the e-ray varies between the maximum value along the optic axis and the minimum value in a direction perpendicular to the optic axis. The velocity of the o-ray is the same in all directions and is greater than that for the e-ray.
- (ii) The refractive index μ_0 is less than μ_e except along the optic axis, where the two are, equal.
- (iii) The wavefront for the o-ray lies outside the wavefront for the e-ray.

Examples: Quartz, aragonite, and topaz.

5.8 HUYGENS' THEORY OF DOUBLE REFRACTION

Huygens explained refraction of light in ordinary media in terms of his wave theory of light. He extended it to explain the phenomenon of double refraction exhibited by certain crystals. His explanation is based on the following assumptions:

- (i) When a beam of light strikes a double refracting crystal, each point on the refracting surface acts as the source for two secondary wavefronts which spread out into the crystal.
- (ii) One wavefront obeys the ordinary laws of refraction which travels with the same speed in all directions. Hence, the corresponding wave surface is spheroid and is called the **o-wavefront**. He also assumed that the wave surface corresponding to the other wavefront is an ellipsoid of revolution.
- (iii) The second wavefront travels with different velocities in different directions. Hence, the wave surface is an ellipsoid which is called the **e-wave front**.
- (iv) The rays corresponding to the two wavefronts merge along the optic axis. The ordinary and extraordinary wave velocities are equal along the optic axis.
- (v) The direction of propagation of extraordinary wavefront depends on the direction of the optic axis relative to the refracting surface.

With these assumptions, Huygens could explain the phenomenon of double refraction satisfactorily.

5.9 NICOL PRISM

The phenomenon of double refraction in calcite is used to construct a very convenient and efficient polarizer and analyzer. If one of the linearly polarized light rays from a doubly refracting crystal is eliminated, the emerging light will be a linearly polarized one. In a Nicol prism, the o-ray is eliminated by total internal reflection, so that only the e-ray is transmitted through.

5.9.1 Construction

A calcite crystal, whose length is about three times its breadth, is taken. Any plane that passes through the optic axis and is normal to a face of the crystal is called a **principal section**.

Let $ABCD$ be a principal section of the crystal [Fig. 5.9(a)]. The two end faces AB and CD are cut down by grinding and polishing so that the angles are 68° and 112° in the principal section instead of the usual 71° and 109° . It is then cut into two and the two parts are then cemented together with Canada balsam whose refractive index ($\mu = 1.526$) lies between the refractive indices of calcite for the ordinary ($\mu_o = 1.658$) and extraordinary ($\mu_e = 1.486$) rays [Fig. 5.9(b)].

Action

For the e-ray, Canada balsam is optically denser than calcite. Therefore, the e-ray is refracted through Canada balsam and the crystal. However, for the o-ray Canada balsam is optically rarer. Hence for angles greater than the critical angle it will be totally reflected. The critical angle for the two media is given by

$$\theta = \sin^{-1} \frac{1.526}{1.658} \cong 67^\circ \quad (5.14)$$

If the angle of incidence on the face AB is limited to very small values ($< 14^\circ$), the o-ray gets totally reflected at the calcite – Canada balsam interface and is absorbed by lamp black on the side of the prism. Thus, the e-ray alone comes out of the Nicol, which is linearly polarized.

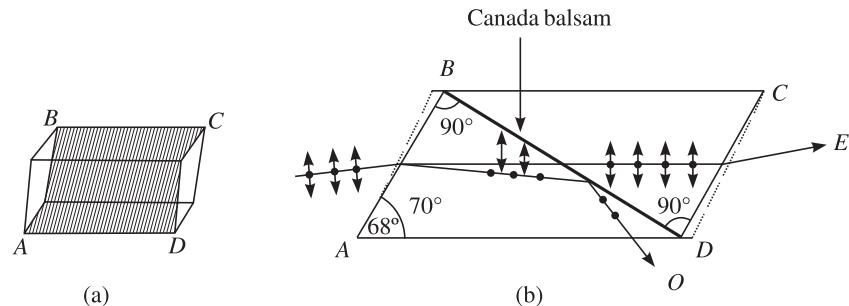


Fig. 5.9 (a) $ABCD$ is a principal section of a calcite crystal, (b) Diagram of a Nicol prism showing the o-ray and e-ray.

5.9.2 As an Analyzer

A Nicol prism can be used as an analyzer of light falling on it. If unpolarized light is allowed to fall on a Nicol prism, and the prism is rotated, the intensity of transmitted light is not affected. If the incident light is partially polarized, the intensity of the transmitted light varies with the rotation of the prism, but never becomes zero. With polarized incident light, on rotation of the prism, the intensity of the transmitted light will be maximum in one position of Nicol and zero when it is further rotated through 90° . A Nicol prism transmits vibrations parallel to its principal plane only. The maximum transmission occurs when the vibrations of the incident light are in the principal plane of the prism, and no transmission occurs when the vibrations are perpendicular to the principal plane. Thus, the Nicol prism can be used as an analyzer.

5.10 RETARDATION PLATES

When light falls normally on a doubly refracting uniaxial crystal plate cut with the optic axis parallel to the refracting surfaces, a phase difference between the ordinary and extraordinary waves is introduced. The crystal is then referred to as a **retardation plate**.

Let the refractive index of the plate for the o-ray be μ_o and for the e-ray be μ_e . If t is the thickness of the plate, the optical path for the o-ray is $\mu_o t$ and for the e-ray is $\mu_e t$. The path difference between the two rays is given as

$$d = (\mu_o - \mu_e)t, \quad \mu_o > \mu_e \quad (5.15)$$

By selecting crystal plates of suitable thickness, we can have devices for producing the required path difference between the o and e-rays. Two such plates are the quarter and half wave plates.

5.11 QUARTER WAVE PLATE

A quarter wave plate is a doubly refracting uniaxial crystal plate that introduces a phase difference of $\pi/2$ (a quarter of a cycle) or a path difference of $\lambda/4$ between the o and e-vibrations when light is incident normally on it. For a negative crystal, the optical path difference of the o and e-vibrations is $(\mu_o - \mu_e)t$, where t is the thickness of the crystal. The correct thickness for a quarter wave plate can be computed as

$$\text{Path difference } \frac{\lambda}{4} = (\mu_o - \mu_e)t \quad (5.16)$$

$$\text{Negative crystal } t = \frac{\lambda}{4(\mu_o - \mu_e)} \quad (5.17)$$

$$\text{Positive crystal } t = \frac{\lambda}{4(\mu_e - \mu_o)} \quad (5.18)$$

In a quarter wave plate, when the incident light is linearly polarized, the emergent beam will be elliptically polarized. However, there are two special cases. First, if the vibration of the incident light makes an angle of 45° to the optic axis, the emergent light will be circularly polarized, and second, if the vibration direction is parallel to or perpendicular to the optic axis, the emergent light is identically polarized.

Quarter wave plates are usually made of quartz or mica. Although mica is biaxial, there is a colourless form for which the angle between the optic axes is vanishingly small.

5.12 HALF WAVE PLATE

A half wave plate introduces a phase difference of π or a path difference of $\lambda/2$ between the o and e-vibrations when light is incident normally on the face of the crystal. It is also made from a doubly refracting uniaxial crystal, cut with its refracting faces parallel to the optic axis. The thickness t of the crystal is given by

$$\text{Negative crystal } t = \frac{\lambda}{2(\mu_o - \mu_e)} \quad (5.19)$$

$$\text{Positive crystal } t = \frac{\lambda}{2(\mu_e - \mu_o)} \quad (5.20)$$

The effect of a half wave plate is to alter the direction of vibration of the linearly polarized light by the angle 2θ , where θ is the angle between the incident vibrations and the optic axis.

5.13 PRODUCTION OF DIFFERENT TYPES OF POLARIZED LIGHT

Various types of polarized light are as follows:

Linearly polarized light

When unpolarized monochromatic light is passed through a Nicol prism, it splits into the ordinary and extraordinary components. Both the components are linearly polarized. The ordinary component is totally internally reflected at the Canada balsam layer, while the extraordinary component is transmitted through the prism. The emergent light is linearly polarized with its vibrations parallel to the shorter diagonal of the end face of the crystal.

Circularly polarized light

The emergent linearly polarized light from a Nicol prism is allowed to fall normally on a quarter wave plate such that the vibration direction is at 45° to the optic axis of the plate. The emergent light from the quarter wave plate is circularly polarized.

Elliptically polarized light

Linearly polarized light is obtained first by passing ordinary monochromatic light through a Nicol prism. The linearly polarized light is then allowed to fall normally on a quarter wave plate such that the vibration direction in the incident light makes an angle θ other than 0° , 45° and 90° with the optic axis of the plate. The light emerging from the quarter wave plate will be elliptically polarized.

5.14 POLARIZATION BY SELECTIVE ABSORPTION

5.14.1 **Dichroism**

Certain crystals have the characteristics of selective absorption. When unpolarized light enters a tourmaline crystal, double refraction takes place and the o-ray is completely absorbed in the crystal. However, the e-ray whose vibration is parallel to the optic axis is transmitted. This phenomenon of selective absorption is sometimes referred to as **dichroism**. Tourmaline crystals, thus take in unpolarized light and transmit only the linearly polarized light. When two such crystals are lined up parallelly, the linearly polarized light from the first passes through the second. If the second is rotated through 90° with respect to the first, the light is completely absorbed.

5.14.2 **Polaroid**

Tourmaline crystals are not used in practice and do not transmit white light as they are yellow in colour. A more satisfactory material which works on this principle is the **polaroid**. It is made in the form of thin films. They are made from the needle shaped crystals of the organic

compound iodosulphate of quinine. These needle shaped crystals are lined up with their optic axes parallel and embedded in a nitrocellulose mastic (gum-resin). These crystals act like a tourmaline crystal by absorbing one component of polarization and transmitting the other. No light passes through two polaroid films in crossed position.

Another common type of polarizer used today is the *H*-sheet. It has a matrix of parallel chains of the synthetic material polyvinyl alcohol impregnated with iodine. Another common one is the *K*-sheet which is a polyvinylene polarizer similar to *H*-sheet, but more resistant to high temperatures and humidity.

Polaroids are used in the study of optical properties of metals, for the control of the intensity of light and for the production of nonglare spectacles, headlights and wind shields of automobiles. Polaroid filters are useful in photography for cutting the unwanted glare from surfaces such as cover glasses of paintings, shining metallic surfaces, etc. Polaroids are extensively used in the production of 3-dimensional movie pictures.

5.15 ANALYSIS OF LIGHT OF UNKNOWN POLARIZATION

By using a polarizer and a quarter wave plate, we can analyse a light of unknown polarization. If light is linearly (plane) polarized and incident on a polarizer (Nicol prism) which is rotated about the optic axis, the transmitted light will vary between a maximum and zero. If the original beam is partially polarized or a mixture of linearly polarized and circularly polarized, the transmitted light will not be extinguished. Hence, to detect circularly polarized light, the given light is first passed through a quarter wave plate which converts circularly polarized light into a linearly polarized light. When this emergent light is viewed through an analyser, it shows variation in intensity with zero minimum on rotation. If the original light is an unpolarized one, it will remain as such when passed through the quarter wave plate. Figure 5.10 shows various steps involved in the analysis of light of unknown polarization using a polarizer and a quarter wave plate.

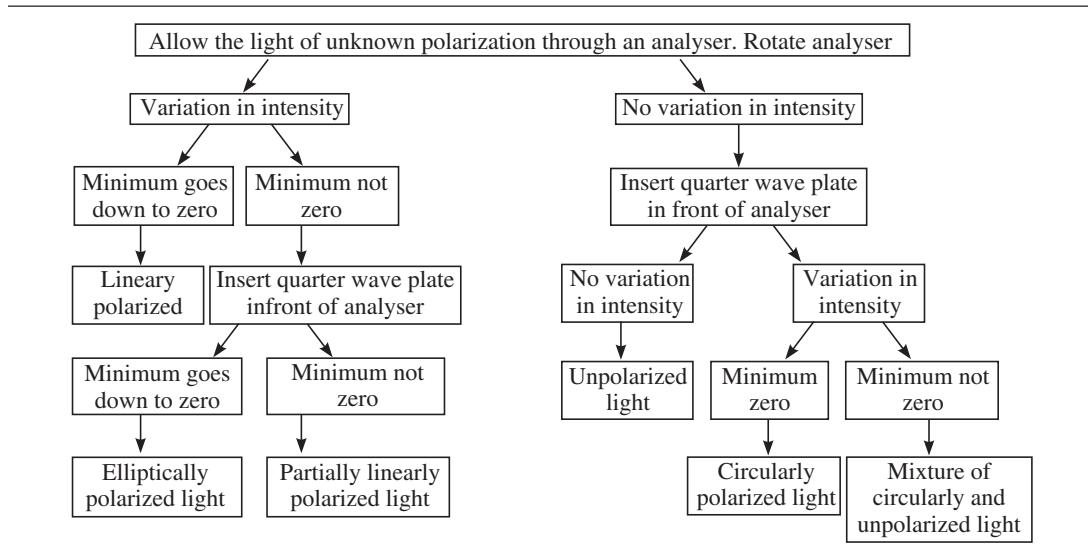


Fig. 5.10 Analysis of light of unknown polarization.

5.16 KERR EFFECT

Kerr electro optic effect is a double refraction induced in solids and liquids by an electric field. The medium acquires precisely the optical character of a uniaxial crystal with the optic axis parallel to the field direction. The double refraction in the medium can be specified in terms of an ordinary and an extraordinary refractive indices. The difference in refractive indices when propagation is at right angles to the field is given by

$$\mu_e - \mu_o = K\lambda E^2 \quad (5.21)$$

where K is the Kerr constant which may be positive or negative, λ is the wavelength, and E is the field strength. The constant K depends on wavelength and temperature.

Kerr cell is a glass cell containing a liquid and two parallel electrodes connected to the positive and negative of a power supply. The liquid is usually nitrobenzene since it has a large Kerr constant. Such a cell inserted between crossed polarizer and analyzer constitutes a useful optical device called **electrooptic shutter**. When the electric field is off, no light is transmitted by the analyzer. When the cell is oriented at 45° , the vibrations from the polarizer are broken up into two equal components parallel and perpendicular to the field. These travel with different speeds, and hence a phase difference is introduced. Consequently, the light emerges as elliptically polarized light.

5.17 FARADAY EFFECT

Faraday in 1845 discovered that when a transparent medium is subjected to a strong magnetic field, it becomes optically active. When linearly polarized light is passed through the medium in a direction parallel to the applied magnetic field, the plane of polarization is rotated. This phenomenon, called **Faraday effect**, is exhibited by a number of solids, liquids and gases. The angle of rotation θ is given by

$$\theta = VBl \quad (5.22)$$

where B is the magnetic induction, l is the length of the path in the field and V is a constant called **Verdet's constant**. V is different for different substances. A positive value for V indicates that the rotation is in the clockwise direction when looking along the direction of the field.

5.18 PHOTOELASTICITY

Photoelasticity is a technique widely used in the engineering world for the analysis of stress distribution in mechanical and structural components using polarized light.

5.18.1 Basic Ideas

Isotropic substances when subjected to mechanical stress, become temporarily double refracting or **birefringent**. When two polarizers are placed in crossed position, no light comes out of the second one, which acts as the analyser. When the specimen without stress is placed between the polarizer and analyser, there will not be any change in the output of the analyser. However, when stress is applied to the specimen, it becomes optically active. The linearly polarized light

passing through the specimen under stress becomes elliptically polarized, and as a result the output of the analyser result to changes in the pattern. A study of the pattern helps to understand the nature of the stress applied to the specimen. To investigate the stress distribution in complex engineering structures, transparent replica of these structures are made with materials like bakelite or zylonite, and placed between the crossed polarizers. An alternate method is to apply a plastic coating to the real structure and examine it in reflected light.

5.19 INDUCED BIREFRINGENCE

Materials like glass, plastic, cellulose, bakelite, etc., are not doubly refracting and are called **optically isotropic**. If they are subjected to mechanical stress, the material acquires double refraction property. This phenomenon of induced optical anisotropy in isotropic material under the action of external forces is referred to as **induced birefringence** or **artificial double refraction**. The induced anisotropy disappears on the removal of the external forces.

Under stress, the light passing through the material splits into ordinary and extraordinary rays. Then the difference between the refractive indices,

$$|\Delta\mu| = \mu_e - \mu_o \quad (5.23)$$

will be a measure of the induced birefringence. Experimentally, it has been found that the induced birefringence at any point inside the material is directly proportional to the stress σ experienced by the material at that point. That is

$$\Delta\mu \propto \sigma \quad \text{or} \quad \Delta\mu = c\sigma \quad (5.24)$$

The constant c is called the **stress-optic coefficient**. Induced birefringence is the basis of photoelastic stress analysis, a method widely used in the engineering world.

Crystalline birefringence occurs if a crystal has an asymmetric structure. All anisotropic crystals and many biological materials such as collagen, cellulose and muscle fibers are some of the examples. These biological materials are so densely packed that liquid cannot penetrate them.

5.20 STRESS ON A PHOTOELASTIC MATERIAL—AN ANALYTICAL TREATMENT

We shall now consider an analytical treatment of the effect of stress on an isotropic body placed between a polarizer P and an analyser A [Fig. 5.11(a)]. Let the linearly polarized light coming from the polarizer is represented by

$$E = A \cos \omega t \quad (5.25)$$

On entering the sample material M under stress, the light splits into two components, one along the fast σ_1 -axis and the other along the slow σ_2 -axis, which is perpendicular to the σ_1 -axis [Fig. 5.11(b)]. The one along the σ_1 -axis makes an angle θ with the optic axis. The two components are:

$$\text{Along the } \sigma_1\text{-axis} \quad E_1 = A \cos \omega t \cos \theta \quad (5.26)$$

$$\text{Along the } \sigma_2\text{-axis} \quad E_2 = A \cos \omega t \sin \theta \quad (5.27)$$

Let E'_1 and E'_2 be the two components of the light vectors leaving the stressed model. Since E'_1 is along the fast axis, there will be a phase change δ between the two.

$$\mathbf{E}'_1 = A \cos(\omega t + \delta) \cos \theta \quad (5.28)$$

$$\mathbf{E}'_2 = A \cos \omega t \sin \theta \quad (5.29)$$

As the analyzer is oriented at right angles to the polarizer the light vectors after transmission through the analyser are [Fig. 5.11(c)].

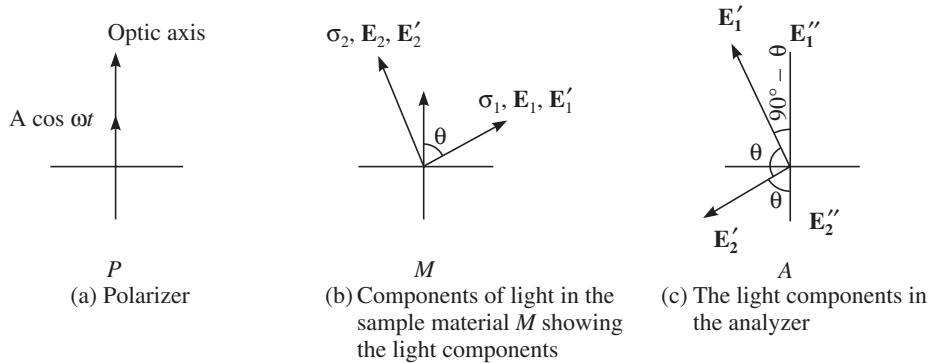


Fig. 5.11 Effect of stress on an isotropic body.

$$\mathbf{E}''_1 = \mathbf{E}'_1 \cos(90^\circ - \theta) = \mathbf{E}'_1 \sin \theta \quad (5.30)$$

$$\mathbf{E}''_2 = \mathbf{E}'_2 \cos \theta \quad (5.31)$$

As the two components \mathbf{E}''_1 and \mathbf{E}''_2 are in opposite directions, the net light vector transmitted through the analyzer is

$$\mathbf{E} = \mathbf{E}''_1 - \mathbf{E}''_2 = \mathbf{E}'_1 \sin \theta - \mathbf{E}'_2 \cos \theta \quad (5.32)$$

Substituting the values of \mathbf{E}'_1 and \mathbf{E}'_2

$$\begin{aligned} \mathbf{E} &= A \cos(\omega t + \delta) \cos \theta \sin \theta - A \cos \omega t \sin \theta \cos \theta \\ &= A \cos \theta \sin \theta [\cos(\omega t + \delta) - \cos \omega t] \\ &= -A \sin 2\theta \sin\left(\omega t + \frac{\delta}{2}\right) \sin \frac{\delta}{2} \end{aligned} \quad (5.33)$$

Intensity $I \propto E^2$. Hence,

$$\begin{aligned} I &= A^2 \sin^2 2\theta \sin^2\left(\omega t + \frac{\delta}{2}\right) \sin^2 \frac{\delta}{2} \\ &= I_0^2 \sin^2 2\theta \sin^2\left(\omega t + \frac{\delta}{2}\right) \sin^2 \frac{\delta}{2} \end{aligned} \quad (5.34)$$

where $I_0 = A^2$, is the maximum transmitted light intensity. Next, we shall consider two special cases, the isoclinic fringes and isochromatic fringes.

5.20.1 Isoclinic and Isochromatic Fringes

Isoclinic fringes are the interference fringes where one of the principal stress direction coincides with the axis of the polarizer. When $2\theta = n\pi$, $n = 0, 1, 2, \dots$, in Eq. (5.34)

$$\sin^2 2\theta = 0 \quad \text{or} \quad I = 0$$

That is, when one of the principal stress directions coincides with the axis of the polarizer, extinction occurs. The fringe pattern observed is called **isoclinic fringes**, and these are used to determine the principal stress direction in the photoelastic material.

Isochromatics are the interference fringes each of which is the locus of points of equal stress difference. When $\frac{\delta}{2} = n\pi$, $n = 0, 1, 2, \dots$, from Eq. (5.34), we obtain

$$\sin^2\left(\frac{\delta}{2}\right) = 0 \quad \text{or} \quad I = 0$$

That is, when the principal stress difference is either zero ($n = 0$) or sufficient to produce an integral number of wavelengths of retardation ($n = 1, 2, 3, \dots$), extinction occurs. The fringe pattern observed is called **isochromatic fringes**.

5.21 PHOTOELASTIC BENCH

In Section 5.18, we learnt that when an isotropic material is under stress, it becomes doubly refracting. **Photo-elastic bench** is an arrangement used to study the photo-elasticity of materials induced by stress applied. Figure 5.12 illustrates the schematic diagram of such a bench.

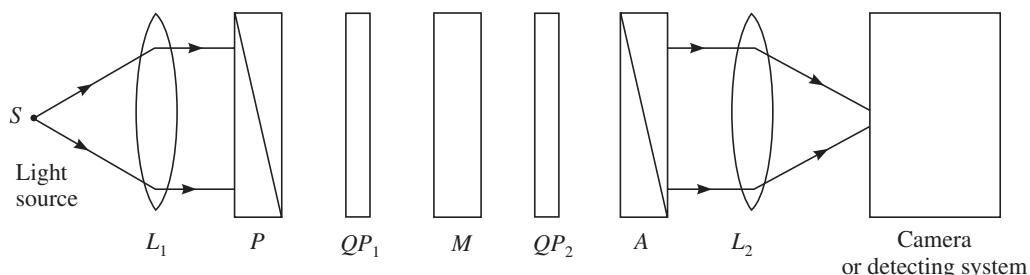


Fig. 5.12 Photoelastic bench. In Fig 5.12, P is polarizer, A is analyzer QP_1 and QP_2 are quarter waveplates and M is the photo-elastic material.

A monochromatic light source is kept at the principal focus of a lens L_1 . The light rendered by L_1 is parallel which falls on a polarizer P . The light coming out of it will be a linearly polarized one. When this falls normally on a quarter wave plate such that the vibration direction is at 45° to the optic axis, the emergent light will be a circularly polarized. M is the photoelastic material kept on a loading frame that has arrangements to apply suitable loads on the material. The light emerging from M enter quarter wave plate again which converts the circularly polarized light into a linearly polarized one. We then have an analyser that analyses the beam coming from the sample under stress. The fringe system emerging from the analyzer then enters a camera or other suitable detecting system.

Dark contours represent regions where stress is not applied. Coloured contours indicate the regions that are under stress. By analysing the pattern of the fringes, we can study the distribution of stress on the material.

SOLVED EXAMPLES

Example 5.1 What is the angle of incidence for complete polarization to occur on reflection at the boundary between water ($\mu = \frac{4}{3}$) and flint glass ($\mu = 1.72$) assuming that the light comes from the side of (i) the water (ii) the glass

Solution: (i) For polarization to occur on flint glass, $\tan i = \frac{\mu_g}{\mu_w} = \frac{1.72}{\frac{4}{3}} = 1.29$

$$\therefore i = 52.2^\circ$$

(ii) For polarization to occur on the water, $\tan i = \frac{\frac{4}{3}}{1.72} = 0.7752$

$$\therefore i = 37.78^\circ$$

Example 5.2 Unpolarized light falls on a Nicol prism. Polarized light emerging from it falls on another crossed Nicol. If the crossed Nicol is rotated through 30° , calculate the percentage of incident light transmitted.

Solution: Let the intensity of the unpolarized light be $2I_0$. Then the intensity of light falling on the second Nicol is I_0 . When the Nicols are crossed, the angle between their principal planes is 90° . If one of them is rotated through 30° , the angle between their planes of transmission is either 60° or 120° . If I is intensity of the emerging light from second Nicol, according to Malus law, then

$$I = I_0 \cos^2 60^\circ = I_0 \cos^2 120^\circ = I_0 \left(\frac{1}{2}\right)^2 = \frac{I_0}{4}$$

$$\text{Percentage transmission} = \frac{I_0}{2I_0} \times 100 = 12.5\%$$

Example 5.3 Calculate the thickness of a quarter wave plate for light of wavelength 6000 \AA . The refractive index for the ordinary and extraordinary rays are 1.54 and 1.55, respectively.

Solution: Since $\mu_e > \mu_o$ the crystal is positive. The thickness of the quarter wave plate is then

$$t = \frac{\lambda}{4(\mu_e - \mu_o)} = \frac{6000 \times 10^{-8} \text{ cm}}{4(1.55 - 1.54)} = 0.0015 \text{ cm}$$

Example 5.4 A calcite plate behaves as a half wave plate for a particular wavelength λ . Assuming the variation of refractive index with λ to be negligible, how would the above plate behave for light of wavelength 2λ .

Solution: The thickness of a half wave plate of calcite for wavelength λ is

$$t = \frac{\lambda}{2(\mu_o - \mu_e)} = \frac{2\lambda}{4(\mu_o - \mu_e)}$$

Hence, the half wave plate for λ will behave as a quarter wave plate for 2λ , if the variation of refractive index with wavelength is negligible.

Example 5.5 The values of μ_o and μ_e for quartz are 1.5418 and 1.5508, respectively. Calculate the phase retardation for $\lambda = 500\text{ nm}$ when the plate thickness is 0.032 mm.

Solution: The path difference between extraordinary and ordinary rays for quartz plate of thickness t is given by

$$\begin{aligned} \text{Path difference} &= (\mu_e - \mu_o)t \\ &= (1.5508 - 1.5418)(32 \times 10^{-6} \text{ m}) = 2.88 \times 10^{-7} \text{ m} \end{aligned}$$

$$\begin{aligned} \text{Phase retardation} &= \frac{2\pi}{\lambda} \times \text{Path difference} \\ &= \frac{2\pi}{500 \times 10^{-9} \text{ m}} (2.88 \times 10^{-7} \text{ m}) = 1.152 \pi \text{ rad} \end{aligned}$$

Example 5.6 The critical angle for total internal reflection at a boundary between two materials is 52° . What is Brewster's angle at this boundary?

Solution: If C is the critical angle

$$\sin C = \mu_2$$

By Brewster's law

$$\mu_2 = \tan i_B$$

where i_B is Brewster's angle

Combining the two equations, we get

$$\tan i_B = \sin C = \sin 52^\circ = 0.788$$

$$i_B = 38^\circ$$

REVIEW QUESTIONS

5.1 What do you understand by polarization of light?

5.2 Define plane of vibration and plane of polarization.

5.3 State and explain Brewster's law.

- 5.4 What do you understand by dichroism?
- 5.5 What is a polaroid? Explain some of its uses.
- 5.6 Explain the phenomenon of double refraction.
- 5.7 With reference to a crystal, explain the terms (i) optic axis (ii) principal section (iii) principal plane.
- 5.8 What are uniaxial and biaxial crystals? Give examples.
- 5.9 What are negative and positive crystals? Give examples.
- 5.10 Distinguish between linearly polarized, circularly polarized and elliptically polarized light.
- 5.11 What is a quarter wave plate? Explain its uses.
- 5.12 Assuming the variation of refractive index with λ to be negligible, show that the half wave plate for λ will behave as a quarter wave plate for 2λ .
- 5.13 How would you show that the ordinary and extraordinary rays coming out of a calcite crystal are polarized with their vibrations at right angles to each other?
- 5.14 Distinguish between circularly polarized and elliptically polarized light.
- 5.15 How will you distinguish circularly polarized light from unpolarized light?
- 5.16 What is Faraday effect?
- 5.17 What is Kerr effect? What is a Kerr cell? Explain its uses.
- 5.18 Explain the principle behind the study of stress applied to an engineering structure using photoelasticity.
- 5.19 What is induced birefringence? Explain the stress-optic law.
- 5.20 What are isoclinic and isochromatic fringes?
- 5.21 Write a brief note on photoelastic bench.
- 5.22 Explain polarization of light by reflection. State Brewster's law. Prove that the angles of incidence and refraction are complementary when complete polarization is obtained by reflection at a plane glass plate.
- 5.23 Explain polarization of light by (i) reflection and (ii) double refraction. Write a note on positive and negative crystals.
- 5.24 What is double refraction? Describe the construction of a Nicol prism. How can it be used as a polarizer and as an analyser?
- 5.25 What is double refraction? Describe a simple experiment to demonstrate double refraction. Write a note on positive and negative crystals.
- 5.26 How can circularly and elliptically polarized light be produced? How would you determine whether a given beam of light is plane polarized, circularly polarized, or elliptically polarized?
- 5.27 What is a quarter wave plate? How would you use it for the production and detection of circularly and elliptically polarized light?

- 5.28** Give an analytical treatment of stress on photoelastic material and explain isoclinic and isochromatic fringes.
- 5.29** Explain photoelasticity. With a block diagram, explain the working of a photoelastic bench.

PROBLEMS

1. A ray of light is incident on a glass plate of refractive index 1.7 at the polarizing angle. Calculate the angle of refraction of the ray.
2. A beam of light travelling in water strikes a glass plate which is also immersed in water. When the angle of incidence is 51° , the reflected beam is found to be linearly polarized. Calculate the refractive index of glass with respect to water.
3. Two Nicols are oriented with their principal planes making an angle of 60° . What percentage of incident unpolarized light will pass through the system?
4. Calculate the thickness of a half wave plate of quartz for light of wavelength 589.3 nm. For quartz $\mu_o = 1.5442$, $\mu_e = 1.5533$.
5. A bundle of unpolarized light passes through two polarizers whose planes of transmission are parallel. Through what angle must one of the polarizers be turned in order to reduce the amount of light to three-fourth of its initial value?
6. Light is reflected at a plane facet of a diamond ($\mu = 2.42$) immersed in oil ($\mu = 1.62$)
 - (i) What is the angle of incidence at which maximum polarization occurs? (ii) What is the angle of refraction?
7. Calculate the minimum thickness of a calcite plate that converts a linearly polarized light into an elliptically polarized light. The refractive indices are $\mu_e = 1.485$ and $\mu_o = 1.656$ at the wavelength 589 nm.
8. A quartz crystal having $\mu_o = 1.5508$ and $\mu_e = 1.5418$ is used as a retardation plate. If the retardation produced is 1.2π and the wavelength of light used is 5461 \AA , what is the thickness of the plate?
9. A quartz crystal having $\mu_o = 1.5508$ and $\mu_e = 1.5418$ is used as a retardation plate. If the wavelength of light used is 5461 \AA and thickness of the plate is 0.025 mm, what is the phase retardation produced?
10. If a half wave plate suitable for wavelength λ is used as a quarter wave plate, what should be the new wavelength? Assume that the variation of refractive index with λ is negligible.

C H A P T E R

6

Crystallography

Since ancient times, solid materials have played an important role in the progress of mankind. Solids can generally be divided into two categories—crystalline and amorphous. In amorphous substances, there is little geometric regularity and periodicity in the way the atoms are arranged in space. By contrast, crystalline substances are characterized by nearly perfect periodicity of atomic structure. In this chapter we shall introduce the fundamentals of crystalline solids.

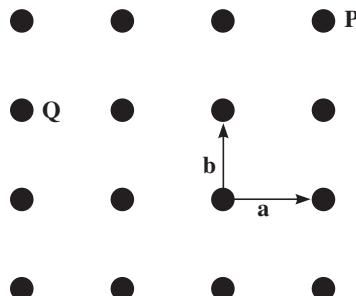
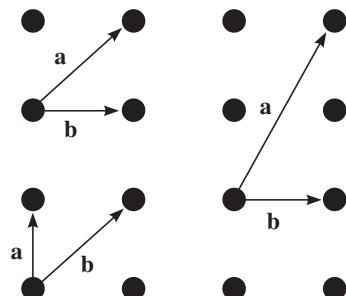
6.1 SPACE LATTICE, BASIS, AND TRANSLATION VECTORS

In crystals, atoms or group of atoms are arranged in a regular pattern. That is, a representative unit is repeated at regular intervals along any direction in the crystal. In non-crystalline solids, the atoms or groups are oriented randomly. We can replace each group by a geometrical point located at the equilibrium position of that group. The resulting periodic arrangements of points in space is known as the **crystal lattice** or simply lattice and the points are referred to as **lattice points**. Such an array of lattice points in two dimensions is called a **plane lattice** and the one in three dimensions, a **space lattice**.

In the plane lattice shown in Fig. 6.1(a), any lattice point can be obtained by repeatedly translating the vectors **a** and **b**. That is, all lattice point are connected by the **translation vector** **T** which is defined by

$$\mathbf{T} = n_1\mathbf{a} + n_2\mathbf{b} \quad (6.1)$$

where n_1 and n_2 are arbitrary integers. The vectors **a** and **b**, called **primitive vectors** or **basis vectors**, are said to generate or span the lattice. Thus, with respect to point *O*, the point **P** = **a** + 2**b** and **Q** = -2**a** + **b**. For any given lattice, the set of primitive vectors is not unique, but many non-equivalent choices are possible as shown in Fig. 6.1(b).

(a) A plane lattice showing the **a**, **b** primitive translation vectors

(b) Different sets of primitive vectors

Fig. 6.1 Primitive vectors in a plane lattice.

It is now essential that we distinguish a lattice from the crystal structure. A lattice is a mathematical entity whereas the crystal structure is a physical one. The repeating physical unit of atom or molecule or radical, etc., that is located at each lattice point is called the **basis** and the crystal structure consists of identical copies of this basis, located at all the points of a lattice. Hence, we may write

$$\text{Lattice} + \text{Basis} = \text{Crystal Structure}$$

Extending to three dimensions, a space lattice can be constructed by the translation operation defined by

$$\mathbf{T} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c} \quad (6.2)$$

where **a**, **b**, and **c** are translation vectors along the three directions of the crystallographic axes and n_1, n_2, n_3 are arbitrary integers. When the operation **T** is applied to any point **r** in the crystal, the resulting point $\mathbf{r}' = \mathbf{r} + \mathbf{T} = \mathbf{r} + n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}$ must be identical in all aspects with the original point **r**. Thus, a **space lattice** can be stated as an infinite 3-dimensional array of points with an arrangement and orientation that appear exactly the same from whichever point the array is viewed.

6.2 BRAVAIS LATTICES

Symmetry operations in a crystal carry the crystal structure into itself. On the basis of symmetry operations, there are 14 ways of arranging the points in space lattices, such that all the lattice points have exactly the same surroundings. The fourteen types of lattices are conveniently grouped into seven **crystal systems** based on the symmetry elements associated with the lattice points. To specify crystal systems conveniently, the axes **a**, **b**, and **c** and angles between axes α , β , and γ are needed. These are illustrated in Fig. 6.2. These 14 point lattices are called **Bravais lattices** and are shown in Fig. 6.3. The seven crystal systems and their unit cell characteristics are given in Table 6.1.

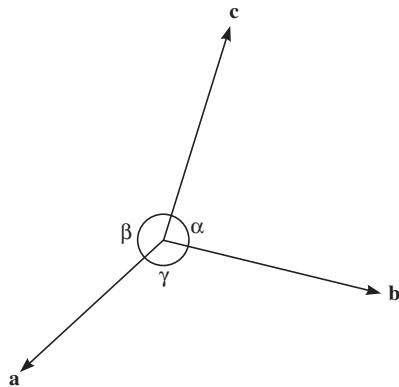


Fig. 6.2 Crystal axes **a**, **b**, **c** and the angles α , β , γ between the axes.

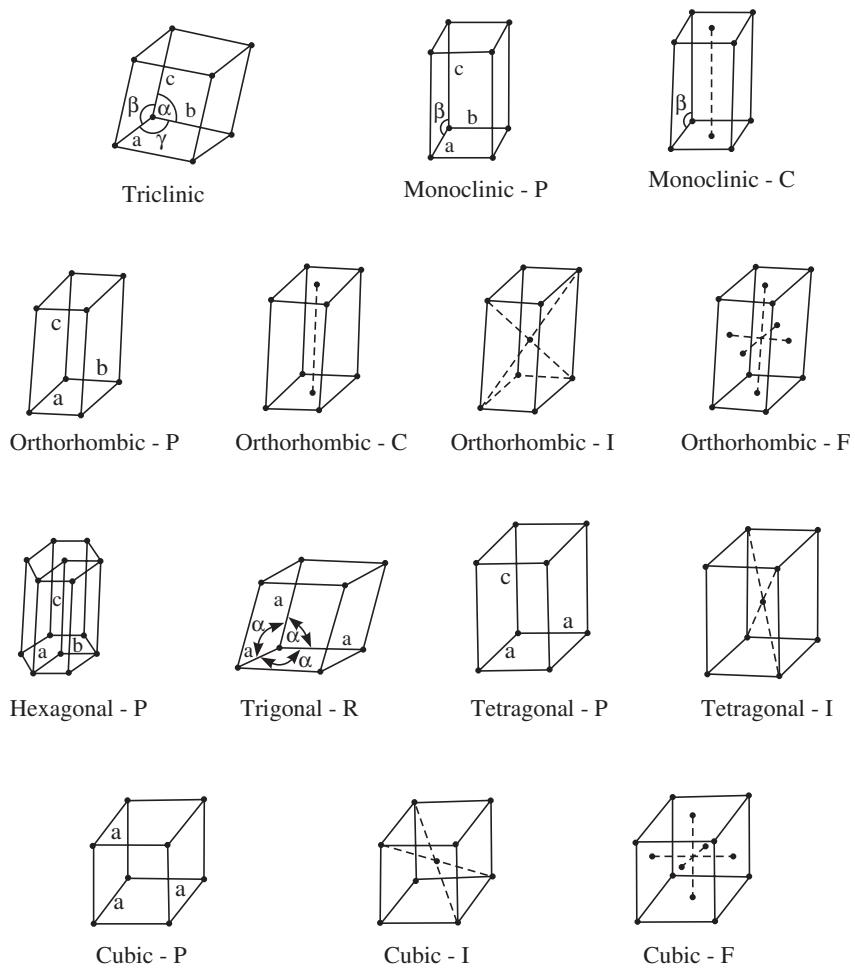


Fig. 6.3 The fourteen Bravais lattices.

Table 6.1 Seven crystal systems, 14 Bravais lattices and conventional unit cell characteristics

Sl. No.	Crystal system	Bravais lattice	Lattice symbol	Unit cell characteristics
1.	Triclinic	Simple (primitive)	P	$\begin{cases} a \neq b \neq c, \\ \alpha \neq \beta \neq \gamma \neq 90^\circ \end{cases}$
2.	Monoclinic	Simple Base centered	P C	$\begin{cases} a \neq b \neq c, \\ \alpha = \gamma = 90^\circ \neq \beta \end{cases}$
3.	Orthorhombic	Simple Base centered Body centered Face centered	P C I F	$\begin{cases} a \neq b \neq c, \\ \alpha = \beta = \gamma = 90^\circ \end{cases}$
4.	Tetragonal	Simple Body centered	P I	$\begin{cases} a = b \neq c, \\ \alpha = \beta = \gamma = 90^\circ \end{cases}$
5.	Cubic	Simple Body centered Face centered	P I F	$\begin{cases} a = b = c, \\ \alpha = \beta = \gamma = 90^\circ \end{cases}$
6.	Hexagonal	Simple	P	$\begin{cases} a = b \neq c, \\ \alpha = \beta = 90^\circ, \gamma = 120^\circ \end{cases}$
7.	Trigonal (Rhombohedral)	Simple	R	$\begin{cases} a = b = c, \\ \alpha = \beta = \gamma < 120^\circ, \neq 90^\circ \end{cases}$

6.3 POINT AND SPACE GROUPS

Translations are not the only type of symmetry operations in a crystal. Apart from translations we can have rotations, reflections, inversions and combinations of these. A symmetry operation that does not involve translation is called a **point operation**. A **point group** is a self consistent set of symmetry elements operating around a point. In crystals, consistent with the translational periodicity, there are only thirty two point groups. There may be other symmetry elements consisting of combinations of point and translational operations. The collection of all the symmetry elements of a structure is called the **space group** and it determines the symmetry of a crystal structure completely. There are 7 crystal systems, 14 Bravais lattices, 32 point groups and 230 space groups. All crystal structures belong to any one of the space groups.

6.4 UNIT CELL

6.4.1 Primitive Unit Cell

A primitive unit cell or simply a primitive cell is the volume of space defined by the three vectors **a**, **b**, and **c**, which when translated by an integral multiple of these vectors just reproduces the entire crystal. A primitive cell is often referred to as **unit cell**. It is a minimum volume cell

which contains only one lattice point. There is no unique way of selecting a primitive cell for a given lattice. An obvious primitive cell is the parallelopiped spanned by the three vectors **a**, **b**, and **c**. Its volume is given by

$$V = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \quad (6.3)$$

Simple cubic lattice

The primitive cell is a cube of side **a**. Hence, the volume of the primitive cell is a^3 and the lattice points per unit volume is 1.

Body centered cubic lattice (bcc)

The primitive cell of a bcc lattice is a rhombohedron. In terms of the cube edge a the primitive translation vectors are:

$$\mathbf{a}' = \frac{a}{2}(\hat{\mathbf{i}} + \hat{\mathbf{j}} - \hat{\mathbf{k}}); \quad \mathbf{b}' = \frac{a}{2}(-\hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}}); \quad \mathbf{c}' = \frac{a}{2}(\hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}})$$

where $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ are unit vectors along the cube edges. The volume of the primitive cell

$$\begin{aligned} V &= (\mathbf{a}' \times \mathbf{b}') \cdot \mathbf{c}' \\ &= \frac{a^2}{4}(\hat{\mathbf{i}} + \hat{\mathbf{j}} - \hat{\mathbf{k}}) \times (-\hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}}) \cdot \mathbf{c}' = \frac{a^2}{4}(2\hat{\mathbf{i}} + 2\hat{\mathbf{k}} \cdot \hat{\mathbf{c}}') \\ &= \frac{a^3}{4}(\hat{\mathbf{i}} + \hat{\mathbf{k}}) \cdot (\hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}}) = \frac{a^3}{2} \end{aligned} \quad (6.3a)$$

Face centered cubic lattice (fcc)

The primitive cell of the fcc lattice is also a rhombohedron. In terms of cube edge a , its primitive translation vectors are:

$$\mathbf{a}' = \frac{a}{2}(\hat{\mathbf{i}} + \hat{\mathbf{j}}); \quad \mathbf{b}' = \frac{a}{2}(\hat{\mathbf{j}} + \hat{\mathbf{k}}); \quad \mathbf{c}' = \frac{a}{2}(\hat{\mathbf{k}} + \hat{\mathbf{i}})$$

The volume of the primitive cell is given by

$$V = \frac{a^2}{4}(\hat{\mathbf{i}} + \hat{\mathbf{j}}) \times (\hat{\mathbf{j}} + \hat{\mathbf{k}}) \cdot \mathbf{c}' = \frac{a^2}{4}(\hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}}) \cdot \frac{a}{2}(\hat{\mathbf{i}} + \hat{\mathbf{k}}) = \frac{a^3}{4}$$

6.4.2 Non-primitive Unit Cell

Instead of using a primitive unit cell, one can span the crystal with a non-primitive unit cell or conventional unit cell. The conventional unit cell is generally selected to be bigger than the primitive cell. It may contain more than one lattice point. Usually, the conventional unit cell

is selected as a cube of edge a for simple cubic, bcc, and fcc lattices. The volume will be a^3 for all the three.

6.5 LATTICE POINTS PER UNIT CELL IN CUBIC LATTICES

Among the Bravais lattices, cubic is the simplest. We shall now calculate the number of lattice points (atoms) contained in a unit cell of the three cubic Bravais lattices.

Simple cubic (sc)

In this, there is one lattice point at each of the corners of the unit cell (see Fig. 6.3). A lattice point at each corner is shared by eight unit cells. Hence, each corner contributes, on the average, $1/8$ th lattice point to the unit cell. Therefore, a simple cubic unit cell should contain $(1/8) \times 8 = 1$ lattice point. Hence, it is a primitive unit cell.

Body centered cubic (bcc)

In body centered cubic, in addition to the lattice points at the corners, there is a lattice point at the body centre. Therefore, the total number of lattice points per unit cell = $1 + 1 = 2$.

Face centered cubic (fcc)

In addition to the lattice points at the corners, there are lattice points occupying the six faces (Fig. 6.3). Hence, the total number of lattice points per unit cell of fcc systems is $(1/8) \times 8 + (1/2) \times 6 = 4$ lattice points.

6.5.1 Density and Number of Atoms per Unit Cell

Let a be the lattice parameter of a cubic lattice, n be the number of atoms per unit cell and ρ be the density of the material. Let M be the atomic or molecular weight.

$$\text{Mass of } n \text{ atoms or molecules} = \frac{M}{N} n$$

where N is the Avogadro's number which is 6.02×10^{26} atoms/k mol. Volume of unit cell is a^3 .

$$\text{Density } \rho = \frac{(M/N)n}{a^3} = \frac{Mn}{Na^3} \quad (6.4)$$

6.6 COORDINATION NUMBER

The strength of packing of atoms in any given structure can be inferred from the number of nearest neighbours that surround each atom. This number is referred to as the **coordination number**. It is defined by the number of equidistant nearest neighbours that an atom has in a given crystal structure. It varies from a minimum of 4 in diamond and zinc blend to a maximum of 12 in close packed structures.

Simple cubic

Consider any corner atom. It has two nearest neighbours each along the x , y and z directions. Thus, the coordination number of a corner atom is 6.

Body centered cubic

For the body centered atom, the atoms at the eight corners are its nearest neighbours, giving the co-ordination number of this atom as eight.

Face centered cubic

Consider the atom at the centre of a face. This face is common to two unit cells. Perpendicular to this common face, there are four perpendicular faces in one unit cell and another four perpendicular faces in the second unit cell. Eight atoms will be there at the centers of eight faces.

All of them are situated at a distance of $\frac{a}{\sqrt{2}}$ from the one under consideration. These eight +

four at the corners of the face under consideration gives a total of twelve. There are, thus, twelve atoms surrounding the atom under consideration and each of them is situated at a distance equal to half of the face diagonal. Therefore, the coordination number of a face centered atom is 12.

6.7 NEAREST NEIGHBOUR DISTANCE

This is the distance between the centres of two nearest neighbouring atoms. We now evaluate the nearest neighbour distance for simple, body centered, and face centered cubic structures. For a crystal containing atoms of the same size, the atomic radius r is half the distance between the nearest neighbours. Assuming that the two nearest neighbours touch each other, the nearest neighbour distance is $2r$, the diameter of the atom. In a cubic system, the length of the face diagonal is $\sqrt{2}a$ and that of the body diagonal is $\sqrt{3}a$ [Fig. 6.4(a)].

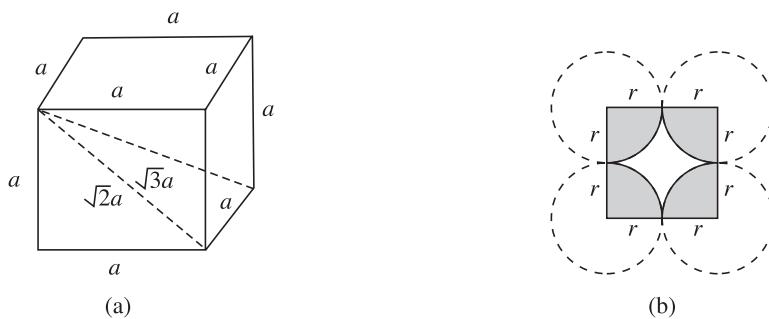


Fig. 6.4 (a) Face diagonal and body diagonal in a cubic system (b) Nearest neighbours in simple cubic

Simple cubic

This structure has atoms only at the corners. Therefore, the atoms touch along cube edges [refer Fig 6.4(b)]. If a is the cube edge and r is the atomic radius, the nearest neighbour distance will be obtained as

$$2r = a \quad \text{or} \quad r = \frac{a}{2} \quad (6.5)$$

$$\text{The nearest neighbour distance} \quad 2r = a \quad (6.5a)$$

Body centered cubic

Since the unit cell is cubic with an atom at each corner and an atom in the centre, the atoms are in contact along body diagonals, Fig. 6.5(a). The length of the body diagonal is $\sqrt{3}a$. Hence,

$$r + 2r + r = \sqrt{3}a \quad \text{or} \quad r = \frac{\sqrt{3}a}{4} \quad (6.6)$$

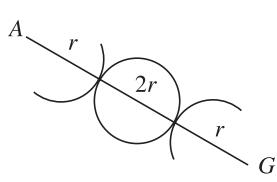
$$\text{The nearest neighbour distance} \quad 2r = \frac{\sqrt{3}a}{2} \quad (6.6a)$$

Face centered cubic

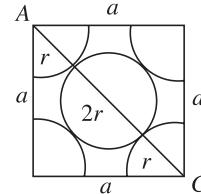
In the close packed face centered cubic structure, atoms occupy each corner and centre of each face as shown in Fig. 6.5(b). The length of the face diagonal is $\sqrt{2}a$. From Fig. 6.5, we have

$$r + 2r + r = \sqrt{2}a \quad \text{or} \quad r = \frac{a}{2\sqrt{2}} \quad (6.7)$$

$$\therefore \text{Neighbour distance} \quad 2r = \frac{a}{\sqrt{2}} \quad (6.7a)$$



(a) Along body diagonal AG
(length $\sqrt{3}a$) in a bcc structure.



(b) Along the face diagonal AC
(length $\sqrt{2}a$) in a fcc structure.

Fig. 6.5 The arrangement of atoms.

6.8 ATOMIC PACKING FACTOR

Packing factor f is defined as the ratio of the volume of atoms occupying the unit cell to the volume of unit cell relating to that structure:

$$f = \frac{\text{Volume occupied by atoms}}{\text{Volume of the unit cell}} \quad (6.8)$$

Simple cubic

The number of atoms in unit cell is one. The atomic radius $r = \frac{a}{2}$, and the volume of the unit cell is a^3 .

$$f = \frac{4}{3}\pi\left(\frac{a}{2}\right)^3 \cdot \frac{1}{a^3} = \frac{\pi}{6} = 0.524 \quad (6.9)$$

Body centered cubic

The number of atoms per unit cell is 2. The atomic radius $r = \sqrt{3} a/4$. Hence,

$$\begin{aligned} \text{Volume of atoms occupying the unit cell} &= 2 \times \frac{4}{3}\pi\left(\frac{\sqrt{3}}{4}a\right)^3 = \frac{\sqrt{3}}{8}\pi a^3 \\ \text{Packing factor } f &= \frac{\sqrt{3}\pi a^3}{8 \times a^3} = 0.68 \end{aligned} \quad (6.10)$$

Face centered cubic

The number of atoms per unit cell is 4 and the atomic radius $r = \frac{a}{2\sqrt{2}}$

$$\begin{aligned} \text{Volume of atoms occupying the unit cell} &= 4 \times \frac{4}{3}\pi\left(\frac{a}{2\sqrt{2}}\right)^3 = \frac{\pi a^3}{3\sqrt{2}} \\ \text{Packing factor } f &= \frac{\pi a^3}{3\sqrt{2} \times a^3} = 0.74 \end{aligned} \quad (6.11)$$

It is obvious that atoms are more closely packed in the face centered cubic as compared to the simple and body centered cubic structures. Table 6.2 summarises different parameters of sc, bcc, and fcc structures.

Table 6.2 Different parameters of sc, bcc, and fcc structures

Parameter	Simple cubic (sc)	Body centered cubic (bcc)	Face centered cubic (fcc)
Volume (unit cell)	a^3	a^3	a^3
Volume (primitive unit cell)	a^3	$\frac{a^3}{2}$	$\frac{a^3}{4}$
Lattice points	1	2	4

(contd.)

Table 6.2 Different parameters of sc, bcc, and fcc structures (contd.)

Parameter	Simple cubic (sc)	Body centered cubic (bcc)	Face centered cubic (fcc)
Coordination number	6	8	12
Atomic radius	$\frac{a}{2}$	$\frac{a\sqrt{3}}{4}$	$\frac{a}{2\sqrt{2}}$
Nearest neighbour distance	a	$\frac{a\sqrt{3}}{2}$	$\frac{a}{\sqrt{2}}$
Packing factor	0.524	0.68	0.74

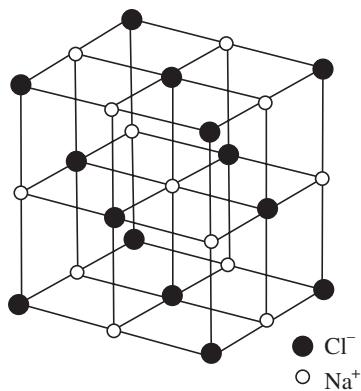
6.9 STRUCTURE OF SODIUM CHLORIDE

Sodium chloride consists of equal numbers of sodium and chlorine ions placed at alternate points of a simple cubic lattice. Each Na^+ ion is surrounded by 6 Cl^- ions and vice versa (Fig. 6.6). Therefore, the coordination number is six. The structure can be considered as two face centered cubic sublattices, one of Cl^- ions having its origin at the point $(0, 0, 0)$ and the other of Na^+ ions having its origin midway along a cube edge, say at point $(a/2, 0, 0)$. The space lattice, is therefore, fcc, with a basis consisting of a Cl^- ion at zero and a Na^+ ion at the centre of the cubic cell, $(a/2, a/2, a/2)$. There are four units of NaCl in a unit cube with ions in positions.

$$\text{Na}^+ : \frac{1}{2}, \frac{1}{2}, \frac{1}{2}; \quad 0, 0, \frac{1}{2}; \quad 0, \frac{1}{2}, 0; \quad \frac{1}{2}, 0, 0$$

$$\text{Cl}^- : 0, 0, 0; \quad \frac{1}{2}, \frac{1}{2}, 0; \quad \frac{1}{2}, 0, \frac{1}{2}; \quad 0, \frac{1}{2}, \frac{1}{2}$$

Some of the crystals having the NaCl arrangements are— KCl , KBr , AgBr , and LiH .

**Fig. 6.6** Structure of sodium chloride.

6.10 CRYSTAL PLANES AND DIRECTIONS

In a crystal with regular periodic lattice, it is necessary to refer to systems of planes within a crystal. The dashed lines in Fig. 6.7(a) represent the intersection of a set of planes passing through the lattice points with the plane of the paper. The perpendicular distance between adjacent planes is called the **interplanar spacing**. The continuous lines in Fig. 6.7(b) represent another set of planes with a different interplanar spacing. Many other such sets of planes with different interplanar spacing can be defined.

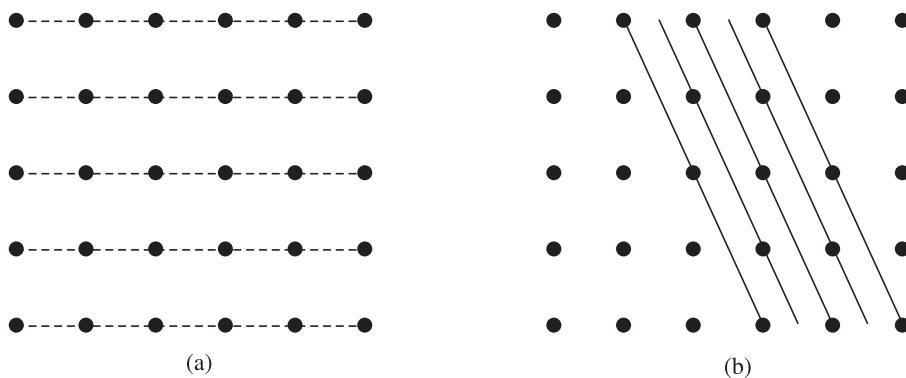


Fig. 6.7 (a) One set of crystal planes (b) Another set of planes with a different interplanar spacing.

Each set of planes has a particular orientation that may be determined by any three points of a plane, provided that the points are not collinear. If each of the points lies on a crystal axis, the plane may be specified by giving the positions of the points along the axes in terms of the lattice constants. For example, if the three points determining the plane have co-ordinates (3,0,0), (0,1,0), (0,0,2) relative to the basis vectors, the plane may be specified by the three numbers 3,1,2. Often, it is more convenient to specify the orientation of a plane by a general set of three numbers called **Miller indices**. (Section 6.11).

Like planes, often we need to specify directions. The indices of a direction are simply the vector components of the direction resolved along the three co-ordinate axes and reduced to smallest integers. It is customary to enclose the indices in square brackets. Thus, the positive a axis has indices [100] and the negative a axis [$\bar{1}00$]. The bar over the digit indicates negative index. The positive b axis has indices [010], negative b axis has indices [0 $\bar{1}$ 0], and so on. The diagonal of the ab -face has indices [110] and that of the ac -face has indices [101]. Accordingly, [111] indicates a body diagonal of a cube. In general, the direction of the vector $ha + kb + lc$ is represented by $(hk\ell)$.

6.11 MILLER INDICES

As indicated, the orientation of crystal planes is specified by a set of three numbers. They may be determined as follows:

- (i) Take any atom in the crystal as the origin. Set co-ordinate axes from this origin in the directions of the basis vectors.

- (ii) Find the intercepts of the plane belonging to the system. Then express the intercepts in the form pa , qb , and rc , where p , q , and r may be fractions.
- (iii) Take the reciprocals of the fractional triplet p , q , r and reduce to the smallest triad of integers h , k , l having the same ratio. This set is called the Miller indices of that system of planes and is denoted as (hkl) .

As an example, consider the plane shown in Fig. 6.8. The intercepts are twice the lattice distance a , four times the lattice distance b , and three times the lattice distance c . The reciprocal of these numbers is $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{3}$. To reduce it to the smallest possible triad of integers multiply each of the reciprocals by 12, which is the L.C.M. of 2, 4, and 3. The resulting triad is 6, 3, 4. The Miller indices for the plane is (634).

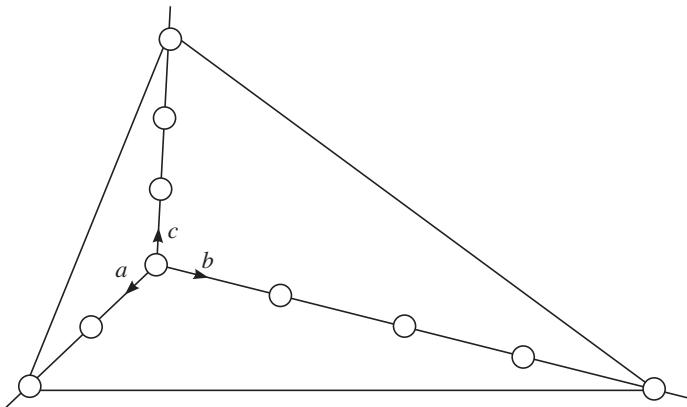


Fig. 6.8 A (634) lattice plane in a crystal lattice.

6.11.1 Significant Features of Miller Indices

- (i) Parallel planes have the same Miller indices. Hence, Miller indices define a set of planes.
- (ii) If a plane cuts an axis on the negative side of the origin, the corresponding index is negative. This is indicated by placing a bar above the index.
- (iii) A parallel plane to one co-ordinate axis has intercept at ∞ . Then the corresponding index is zero.
- (iv) Planes whose Miller indices differ by a permutation of numbers or of minus signs are all crystallographically equivalent. That is, the planes (hkl) , (khl) , (lkh) , $(\bar{h}\bar{k}l)$, $(h\bar{k}\bar{l})$ etc., are all equivalent.

6.12 INTERPLANAR SPACING

Consider a set of planes (hkl) . Let ABC be a plane in that. Take any lattice point O in the plane which is nearest to plane ABC as origin and erect co-ordinate axes in the directions as shown

in Fig. 6.9. Then $OA = \frac{a}{h}$, $OB = \frac{b}{k}$ and $OC = \frac{c}{l}$. Draw a normal OP from the origin to the plane. The distance $OP = d$ is the interplanar spacing. Let this normal OP makes angles α , β and γ with a , b and c axes, respectively. In an orthogonal lattice these directions coincide with the x , y and z axes.

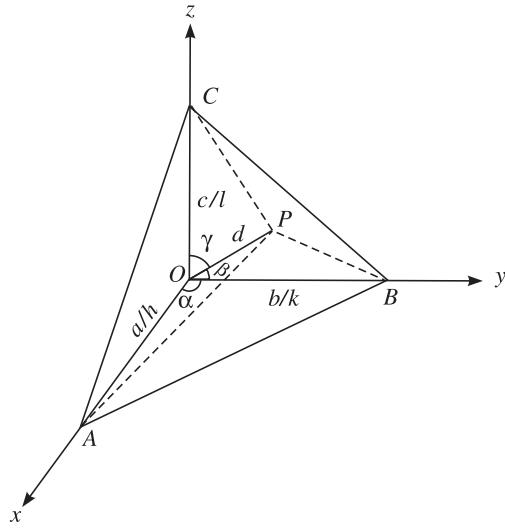


Fig. 6.9 Plane ABC with Miller indices (hkl) along with the direction angles of the normal OP .

$$\text{In triangle } OPA, \cos \alpha = \frac{OP}{OA} = \frac{d}{\frac{a}{h}} = \frac{dh}{a} \quad (6.12)$$

$$\text{In triangle } OPB, \cos \beta = \frac{OP}{OB} = \frac{d}{\frac{b}{k}} = \frac{dk}{b} \quad (6.13)$$

$$\text{In triangle } OPC, \cos \gamma = \frac{OP}{OC} = \frac{d}{\frac{c}{l}} = \frac{dl}{c} \quad (6.14)$$

For an orthogonal co-ordinate system, the law of direction cosines gives

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1 \quad (6.15)$$

Substituting for $\cos \alpha$, $\cos \beta$ and $\cos \gamma$,

$$d^2 \left[\frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2} \right] = 1$$

$$d_{hkl} = \frac{1}{\left[\left(\frac{h}{a} \right)^2 + \left(\frac{k}{b} \right)^2 + \left(\frac{l}{c} \right)^2 \right]^{\frac{1}{2}}} \quad (6.16)$$

This expression is valid only for an orthogonal lattice, that is for cubic, tetragonal, and orthorhombic systems.

For a cubic system

$$a = b = c$$

$$d_{hkl} = \frac{a}{\left[h^2 + k^2 + l^2 \right]^{\frac{1}{2}}} \quad (6.17)$$

For a tetragonal system

$$a = b \neq c$$

$$d_{hkl} = \frac{1}{\left[\left(\frac{h^2 + k^2}{a^2} \right) + \left(\frac{l}{c} \right)^2 \right]^{\frac{1}{2}}} \quad (6.18)$$

For an orthorhombic system $a \neq b \neq c$. Hence, the expression for d is the same as Eq.(6.16).

6.13 X-RAY DIFFRACTION

Max von Laue was the first to suggest in 1912, that a crystalline solid may act on x-rays in a way similar to a grating on visible light. A crystalline solid consists of a regular array of atoms or group of atoms. The crystal is built by the repetition of the fundamental unit, called **unit cell**, which repeats itself throughout the array. Diffraction of x-rays by such units may be visualized as reflections from the different lattice planes of the crystal. Thus, a crystalline solid acts as a three-dimensional grating for x-rays. Indeed, when a collimated beam of monochromatic x-rays is allowed to fall on a crystal, it generates a diffraction pattern.

6.13.1 Bragg's Law

Consider a beam of monochromatic x-rays incident on a crystal, making an angle θ with the surface. Some of the incident rays are reflected at plane 1, some others penetrate to planes 2, 3, 4, ... (Fig. 6.10). It may be noted that angles in x-ray crystallography are measured from the surface and not from the normal to the surface as in the other areas of optics. Let the interplaner spacing of the layers be d . The path difference p between the rays reflected from layer 1 and layer 2 is given by

$$p = AB + BC = d \sin \theta + d \sin \theta = 2d \sin \theta \quad (6.19)$$

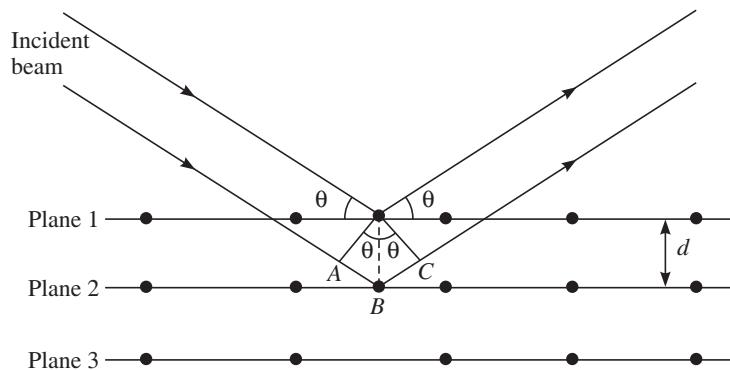


Fig. 6.10 Bragg's reflections from a set of planes repeated by a distance d .

See Section 6.12. For constructive interference of the waves reflected from planes 1 and 2, the path difference must be an integral multiple of a wavelength:

$$2d \sin \theta = n\lambda, \quad n = 1, 2, \dots \quad (6.20)$$

where n is the order number of the intensity maximum. Equation (6.20) is known as **Bragg's Law**. The angle θ will be different for the different orders from the same set of planes. The same relationship holds if the x-rays are transmitted through the crystal.

In what we considered, the lattice planes are parallel to the surface. But there are many other possible sets of planes (Fig. 6.7). The number of scatterers per unit area in these sets of planes will be different. Hence, the corresponding maxima will be of different sizes.

6.13.2 X-ray Spectrometer

Figure 6.11 gives the schematic diagram of the spectrometer developed by Bragg. It is similar to the optical spectrometer. X-rays are collimated and allowed to fall on a crystal mounted on a table at the center. The table is mounted on a graduated scale. Bragg used an ionization chamber as detector. The incident x-ray beam after diffraction from the crystal enters the ionization chamber where its intensity is measured.

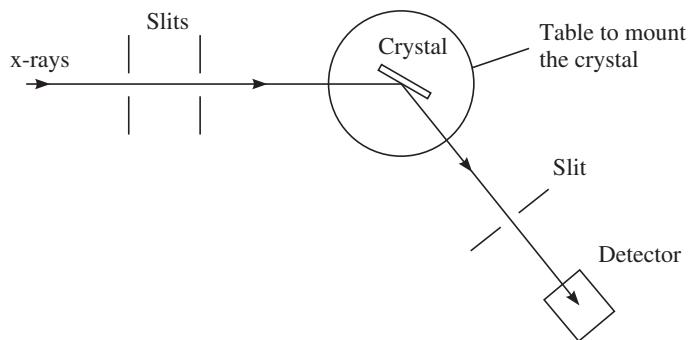


Fig. 6.11 Bragg's x-ray spectrometer.

The modern x-ray spectrometers consist of a highly sensitive electronic detector, display, and recording system. Arrangements for data processing are also available. If the sample is rotated, the angle θ changes and a series of bright and dark regions corresponding to constructive and destructive interference would be observed.

6.13.3 Unit Cell Dimensions

From the analysis of the diffraction pattern, we can obtain the value of θ . Substituting the value of θ in Eq. (6.20), the interplaner spacing d can be calculated. Figure 6.12 shows how the interplaner spacing d is related to the unit cell dimensions a_0 in a cubic case. For the particular family of planes shown in Figure 6.12, by Pythagoras theorem.

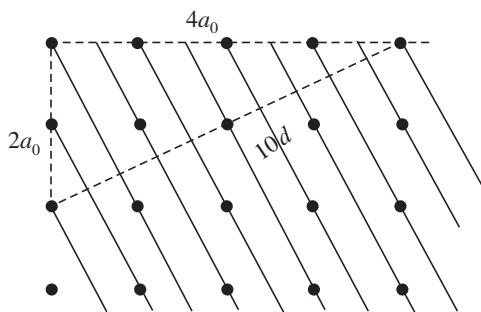


Fig. 6.12 Illustration showing cell dimensions and interplanar spacing.

$$\sqrt{20} a_0 = 10d \quad \text{or} \quad a_0 = \sqrt{5} d \quad (6.21)$$

Thus, unit cell dimension can be evaluated from the interplanar spacing. X-ray diffraction is a very powerful tool for studying the arrangement of atoms and the structure of the unit cell in crystals.

6.14 X-RAY DIFFRACTION—LAUE'S METHOD

To satisfy the Bragg's law $2d \sin \theta = n\lambda$, it is necessary to provide a continuous range of values of either λ or θ . In the Bragg's method, the crystal is rotated, so that the condition is satisfied. In the Laue method, a single crystal is held stationary in a x-ray beam of continuous wavelengths. The crystal selects and diffracts the value of λ for which planes separated by spacing d exist.

The experimental arrangement is shown in Fig. 6.13. X-rays of a range of wavelengths (0.2 \AA to 2 \AA) are allowed to pass through a slit to make it narrow. The collimated beam then falls on the crystal which is held stationary. Each plane in the crystal selects from the incident beam a particular wavelength satisfying the Bragg condition. If monochromatic x-rays are used, a very few sets of planes will satisfy the Bragg condition. A flat photographic film is used the detector. The diffraction pattern consists of a series of spots, called **Laue spots**. The nature of the pattern allows one to determine the symmetry of the crystal. The Laue method is seldom

used for crystal structure determination. However, it is used to orient crystals for solid state experiments and to study the extent of crystalline imperfection under mechanical and thermal treatment.

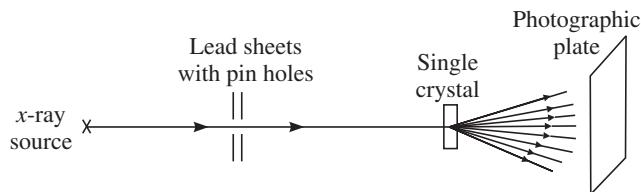


Fig. 6.13 Experimental arrangement for Laue's method.

6.15 X-RAY DIFFRACTION—POWDER METHOD

In the powder method, the incident monochromatic radiation is allowed to strike a powdered specimen or a polycrystalline sample. The experimental arrangement is shown in Fig. 6.14. The powder specimen is taken in a thin capillary tube which is placed at the centre of a cylindrical camera. In the atomic scale, the distribution of the crystallite orientations will be nearly continuous, and therefore is capable of diffracting x-rays. As the crystal axes of the individual grains are randomly oriented, the pattern produced by such a powder would be the same as the one produced by combining the patterns for all possible orientations of a single crystal. For a set of planes satisfying the Bragg condition, the diffracted rays lie on the surface of a cone having semi vertical angle 2θ . Considering all the set of planes, the diffracted rays leave the specimen along the generators of cones concentric with the original beam. The cones intercept the film in a series of concentric rings with a circular spot at the centre corresponding to the undeflected beam (Fig. 6.15).

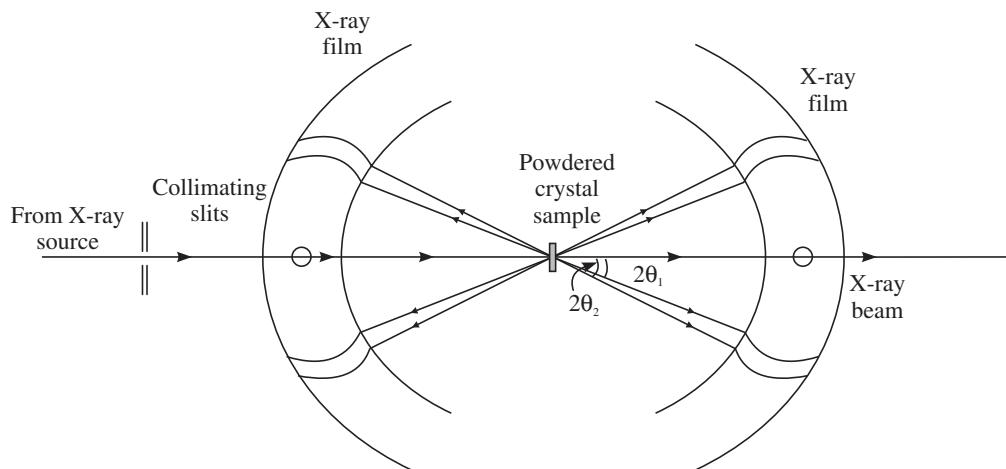


Fig. 6.14 Experimental arrangement for X-ray powder diffraction.

In the diffraction pattern, let s be the distance on the film between the diffraction arcs corresponding to a particular plane (Fig. 6.15). From the figure, the relation between s and the diffraction angle θ is obtained as

$$s = 4\theta R$$

where θ is the Bragg angle in radians and R is the specimen–film distance, usually the radius of the camera. Since the wavelength is known, substitution of θ and λ in the Bragg equation gives d , the interplanar spacing. Using geometric relations connecting d , Miller indices, and unit cell dimensions, we can evaluate the unit cell dimensions.

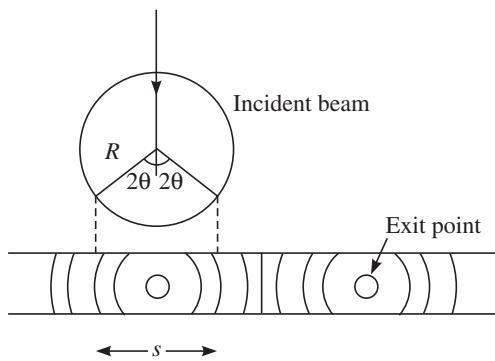


Fig 6.15 A typical powder photograph with the relation between s and R .

SOLVED EXAMPLES

Example 6.1 Calculate the lattice parameter of an NaCl crystal, which has an fcc structure, given that its density is $2.18 \times 10^3 \text{ kg/m}^3$, atomic weight of sodium and chlorine are 23 and 35.5, respectively.

Solution: For an fcc crystal the number of atoms $n = 4$. Molecular weight of NaCl is 58.5. Avogadro's number is 6.02×10^{26} atoms/k mol. From Eq. (6.4), we get

$$\begin{aligned} a^3 &= \frac{Mn}{N\rho} = \frac{(58.5 \text{ kg}) \times 4}{(6.02 \times 10^{26} \text{ atoms/k mol})(2.18 \times 10^3 \text{ kg/m}^3)} \\ &= 178.3 \times 10^{-30} \text{ m}^3 \end{aligned}$$

$$a = 5.63 \times 10^{-10} \text{ m} = 0.563 \text{ nm}$$

Example 6.2 A crystal plane cuts intercepts $3a$, $2b$, and c along the a , b , c axes. Determine the Miller indices of the plane.

Solution: Intercepts: $3a$, $2b$, c

The fractional triplet: 3, 2, 1

Reciprocals: $\frac{1}{3}, \frac{1}{2}, 1$

Multiplying by L.C.M: 2, 3, 6

\therefore Miller indices of the plane (2 3 6)

Example 6.3 Determine the Miller indices of a plane parallel to the b -axis cutting intercepts of $2a$ and $\frac{3c}{2}$ along the x and z axes, respectively.

Solution: Intercepts: $2a, \infty, \frac{3c}{2}$

The fractional triplet: $2, \infty, \frac{3}{2}$

Reciprocals: $\frac{1}{2}, 0, \frac{2}{3}$

Multiplying by LCM: 3, 0, 4.

Miller indices of the plane (3 0 4)

Example 6.4 Sketch the (110) (100) and (111) planes

Solution: (110) plane: Taking the reciprocals of the Miller indices, we get

$$\frac{1}{1}, \frac{1}{1}, \frac{1}{0}$$

Making them whole numbers: 1, 1, ∞

Hence, the intercepts are a , b , and ∞

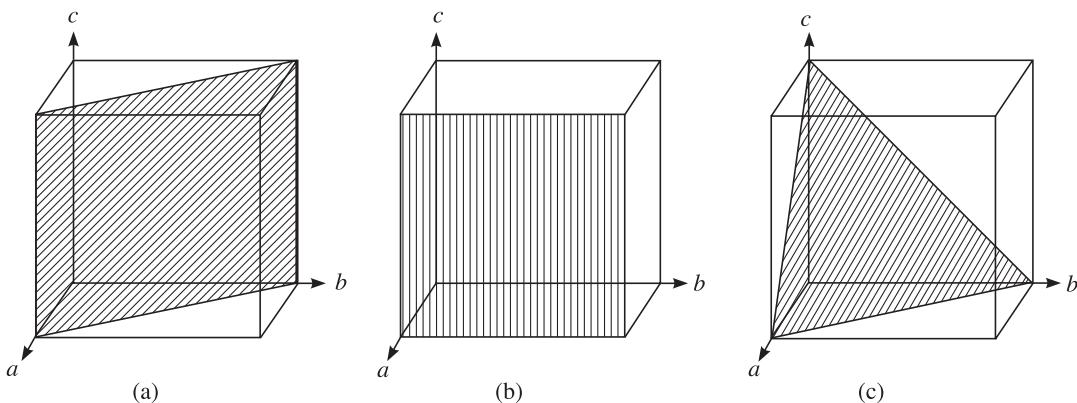


Fig. 6.16 (a) (110) plane, (b) (100) plane, (c) (111) plane.

The (110) plane is the shaded one in Fig. 6.16(a). In the same way, the intercepts for the (100) and (111) planes are a, ∞, ∞ ; and a, b, c , respectively. These planes are represented in Figs. 6.16(b) and (c).

Example 6.5 Calculate the spacing between (i) (110) planes and (ii) (212) planes in a cubic system having the lattice constant $a = 0.43$ nm.

Solution: (i) For (110) planes

$$d_{110} = \frac{a}{(h^2 + k^2 + l^2)^{\frac{1}{2}}} = \frac{0.43 \text{ nm}}{(1^2 + 1^2 + 0^2)^{\frac{1}{2}}} = 0.304 \text{ nm}$$

(ii) For (212) planes

$$d_{212} = \frac{0.43 \text{ nm}}{(2^2 + 1^2 + 2^2)^{\frac{1}{2}}} = \frac{0.43 \text{ nm}}{3} = 0.143 \text{ nm}$$

Example 6.6 Calculate the interplanar spacing of (231) planes of an fcc structure whose atomic radius is 0.175 nm.

Solution: From Eq. (6.7), the atomic radius of an fcc lattice is given by

$$\begin{aligned} r &= \frac{a}{2\sqrt{2}} \quad \text{or} \quad a = 2\sqrt{2}r \\ a &= 2\sqrt{2} \times 0.175 \text{ nm} \\ d_{231} &= \frac{2\sqrt{2} \times 0.175 \text{ nm}}{(2^2 + 3^2 + 1^2)^{\frac{1}{2}}} = 0.132 \text{ nm} \end{aligned}$$

Example 6.7 Find the angles at which 1.44 Å x-rays are diffracted by a polished crystal of rock salt ($d = 2.8$ Å), forming the first and second order maxima.

Solution: Bragg's law states $2d \sin \theta = n\lambda$

$$\text{First order: } \sin \theta = \frac{n\lambda}{2d} = \frac{1.44 \text{ Å}}{2 \times 2.8 \text{ Å}} = 0.2571, \quad \theta = 14.9^\circ$$

$$\text{Second order: } \sin \theta = \frac{2 \times 1.44 \text{ Å}}{2 \times 2.8 \text{ Å}} = 0.5143, \quad \theta = 30.9^\circ$$

Example 6.8 Determine the actual volume occupied by the spheres in an fcc structure as a percentage of the total volume.

Solution: At the 8 corners we have 8 spheres and each is shared by 8 cells. The sphere at the centre of a face is shared by two cells. We have six such faces. Hence, the total number of spheres in a unit cell is given by

$$\left(8 \times \frac{1}{8}\right) + \left(6 \times \frac{1}{2}\right) = 4$$

In an fcc, the atomic radius $r = \frac{a}{2\sqrt{2}}$

$$\frac{\text{Volume of atoms}}{\text{Volume of unit cell}} = \frac{\frac{4}{3} \pi \left(\frac{a}{2\sqrt{2}}\right)^3}{a^3} = \frac{\pi}{3\sqrt{2}}$$

$$\text{Percentage volume occupied} = \frac{\pi}{3\sqrt{2}} \times 100\% = 74.1\%$$

Example 6.9 A beam of x-rays is incident on the (221) plane of a cubic crystal of lattice spacing 2.68 Å. The first order Bragg's reflection is observed at 8.5° glancing angle. Calculate (i) the interplanar spacing (ii) the wavelength of x-rays (iii) the angle at which the second order Bragg reflection occurs.

Solution: (i) For cubic structure the interplanar spacing is given by

$$\begin{aligned} d_{hkl} &= \frac{a}{(h^2 + k^2 + l^2)^{\frac{1}{2}}} = \frac{2.68 \times 10^{-10} \text{ m}}{(2^2 + 2^2 + 1)^{\frac{1}{2}}} \\ &= \frac{2.68 \times 10^{-10} \text{ m}}{3} = 0.893 \times 10^{-10} \text{ m} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad \lambda &= \frac{2d \sin \theta}{n} = 2(0.893 \times 10^{-10} \text{ m}) 0.1478 \\ &= 0.264 \times 10^{-10} \text{ m} = 0.264 \text{ Å} \end{aligned}$$

$$\begin{aligned} \text{(iii)} \quad \sin \theta &= \frac{n\lambda}{2d} = \frac{2(0.264 \times 10^{-10} \text{ m})}{2(0.893 \times 10^{-10} \text{ m})} \\ &= 0.2956 \\ \theta &= 17.2^\circ \end{aligned}$$

Example 6.10 X-rays of wavelength 0.71 Å are diffracted by the (110) plane of a body centered cubic crystal at an angle of 25° in the first order. Calculate the lattice spacing.

Solution: According to Bragg's law

$$d = \frac{n\lambda}{2 \sin \theta} = \frac{0.7 \text{ Å}}{2 \times 0.4226} = 0.8282 \text{ Å}$$

For a cubic system, the lattice spacing

$$a = d (h^2 + k^2 + l^2)^{\frac{1}{2}}$$

$$\begin{aligned}
 &= 0.8282 \text{ \AA} (1^2 + 1^2)^{\frac{1}{2}} \\
 &= 1.17 \text{ \AA}
 \end{aligned}$$

Example 6.11 A beam of x-rays is incident on a crystal with interplanar spacing 0.31 nm. Calculate the wavelength of x-rays, if the first order Bragg reflection takes place at an angle of 9.25°. Also find the maximum order of diffraction possible.

Solution: By Bragg's law

$$\begin{aligned}
 \lambda &= \frac{2d \sin \theta}{n} = 2 \times 0.31 \times 0.1607 \\
 &= 0.0996 \text{ nm}
 \end{aligned}$$

Maximum angle possible is 90°. Hence, the maximum order possible is

$$n = \frac{2d \sin \theta}{\lambda} = \frac{2 \times 0.31 \text{ nm} \times 1}{0.0996 \text{ nm}} = 6.22$$

The maximum order possible is 6.

Example 6.12 A beam of x-rays of wavelength 0.071 nm is diffracted by the (210) plane of a bcc crystal. Find the Bragg reflection angle for the second order diffraction if the interplaner spacing of the (110) plane is 0.195 nm.

Solution: For a bcc system

$$\begin{aligned}
 d_{110} &= \frac{a}{(1^2 + 1^2)^{\frac{1}{2}}} = \frac{a}{\sqrt{2}} \\
 a &= \sqrt{2} \times 0.195 \text{ nm} = 0.2758 \text{ nm} \\
 d_{210} &= \frac{0.2758 \text{ nm}}{(2^2 + 1^2)^{\frac{1}{2}}} = 0.12334 \text{ nm} \\
 \sin \theta &= \frac{2(0.071 \text{ nm})}{2(0.12334 \text{ nm})} = 0.5756 \\
 \theta &= 35.14^\circ
 \end{aligned}$$

Example 6.13 Sodium Chloride has fcc structure. Calculate the distance between the nearest neighbours, if its density is 2182 kg/m³.

Solution: From Eq. (6.4), volume of a unit cell is given by

$$a^3 = \frac{nM}{N\rho}$$

Molecular weight of NaCl = 23 + 35.5 = 58.5

Avogadro's number $N = 6.02 \times 10^{26} / \text{k mol}$

For an fcc structure $n = 4$

$$a^3 = \frac{4 \times 58.5 \text{ kg/k mol}}{(6.02 \times 10^{26} / \text{k mol})(2182 \text{ kg/m}^3)} = 0.178 \times 10^{-27} \text{ m}^3$$

$$a = 0.5627 \times 10^{-9} \text{ m}$$

$$\text{Distance between } \text{Na}^+ \text{ and } \text{Cl}^- = \frac{a}{2} = 0.281 \text{ nm}$$

Example 6.14 The structure of metallic iron changes from bcc to fcc at 910°C and as a result the atomic radii changes from 1.258 Å to 1.292 Å. Calculate the corresponding percentage volume change.

Solution: For a bcc structure, the atomic radius

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

$$a = \frac{4}{\sqrt{3}} \times 1.258 \text{ Å} = 2.905 \text{ Å}$$

$$\text{Volume of unit cell } a^3 = (2.905 \text{ Å})^3 = 24.515 \text{ Å}^3$$

Number of atoms per unit cell = 2

$$\text{Volume occupied by 1 atom} = \frac{24.515}{2} = 12.26 \text{ Å}^3$$

For an fcc structure:

$$a = 2\sqrt{2} r = 2 \times \sqrt{2} \times 1.292 \text{ Å} = 3.654 \text{ Å}$$

$$\text{Volume of unit cell } a^3 = 48.788 \text{ Å}^3$$

Number of atom per unit cell = 4

$$\text{Volume occupied by 1 atom} = \frac{48.788 \text{ Å}^3}{4} = 12.2 \text{ Å}^3$$

$$\begin{aligned} \text{Percentage change in volume} &= \left(\frac{12.26 - 12.2}{12.26} \right) \times 100 \\ &= 0.49\% \end{aligned}$$

REVIEW QUESTIONS

- 6.1** What is a space lattice?
- 6.2** Distinguish between basis vector and basis.
- 6.3** What is a primitive cell? How does the primitive cell of a simple cubic lattice differ from that of a body centered cubic lattice?
- 6.4** What is coordination number? Obtain the coordination number for sc, bcc, and fcc lattices.
- 6.5** List seven crystal systems and their features.
- 6.6** Calculate the packing factor for bcc and fcc lattices.
- 6.7** Write a brief note on the structure of a sodium chloride crystal.
- 6.8** Calculate the number of lattice points per unit cell for sc, bcc, and fcc lattices.
- 6.9** Calculate the atomic radius for bcc and fcc lattices.
- 6.10** What are Miller indices?
- 6.11** Sketch the (220) and (110) planes in a cubic lattice.
- 6.12** Briefly outline the important characteristics of Miller indices.
- 6.13** Define packing factor and lattice constant.
- 6.14** Explain the principle of x-ray diffraction.
- 6.15** How does a crystalline solid behave as a three dimensional grating?
- 6.16** State and explain Bragg's law.
- 6.17** Why is Laue's x-ray diffraction method not as powerful as Bragg's method for structure determination?
- 6.18** Obtain the number of lattice points per unit cell, coordination number, and packing factor with reference to sc, bcc, and fcc lattices.
- 6.19** What are Miller indices? Derive expressions for the interplanar spacing for cubic, tetragonal, and orthorhombic systems.
- 6.20** What is a Bravais lattice? What are the seven crystal systems? Calculate the number of atoms per unit cell with reference to sc, bcc, and fcc lattices.
- 6.21** Explain the principle of x-ray diffraction. Derive Bragg's law. How are interplanar spacing and lattice constant determined using a Bragg spectrometer.
- 6.22** Describe the experimental set up needed for x-ray diffraction by power method. How is it used to determine the unit cell dimensions?

PROBLEMS

1. Determine the Miller indices of the plane that cuts intercepts $2a$, $3b$, and c along the crystallographic axes.
2. Determine the Miller indices of a plane parallel to the c -axis and cutting intercepts of 2 and $3/2$ along the x and y axes, respectively.
3. A crystal plane has Miller indices (221) . Find the intercepts along a , b , and c axes.
4. The atomic radius of silver having fcc structure is 0.152 nm . Calculate the interplanar spacing of (i) (110) planes (ii) (231) planes.
5. Determine the Miller indices of a plane in a cubic crystal of side 0.424 nm which is parallel to the b -axis and cuts intercepts 2 and 1 along the a and c -axes, respectively. Also calculate the interplanar spacing.
6. Calculate the distance between the nearest neighbours in KCl cubic lattice. The density of KCl is 1990 kg/m^3 and its molecular weight is 74.6.
7. When x-rays are incident on the surface of a crystal, the first order diffraction is found to occur at a glancing angle of 9° . Calculate the wavelength of x-ray and the glancing angle for second order diffraction if the spacing between the adjacent planes is 2.51 \AA .
8. X-rays of wavelength 1.4 \AA are found to be Bragg reflected from the (111) plane of an fcc structure. If the lattice parameter of the crystal is 5 \AA , find the angle at which the x-ray is incident on the (111) plane of the crystal.
9. Lithium crystallizes in a bcc structure. If the density of lithium is 530 kg/m^3 , calculate its lattice constant. The atomic weight of lithium is 6.94.
10. In a tetragonal crystal $a = b = 2.52\text{ \AA}$, $c = 1.8\text{ \AA}$. Deduce the value of the interplanar spacing of (111) and (210) planes.
11. The interatomic spacing of a cubic crystal is 2.67 \AA . If x-rays of wavelength 1.54 \AA are diffracted by the (111) planes, what is the diffraction angle in the second order?
12. A beam of x-rays is incident on a cubic crystal having a lattice spacing of 0.282 nm on the plane (100) . The first order Bragg reflection is observed at an angle of 8.5° . At what angle would the second order Bragg reflection occur?
13. Determine the actual volume occupied by the atoms in a body centered cubic structure as a percentage of the total volume.
14. Determine the maximum radius of sphere that can be placed into a body centered cubic structure without affecting the position of other spheres.
15. X-rays of wavelength $\lambda = 1.3\text{ \AA}$ are diffracted in a Bragg spectrograph at an angle of 25° in the first order. What is the interplanar spacing of the crystal?
16. An x-ray beam of a certain wavelength is incident on an NaCl crystal at 30° to a certain family of reflecting planes of spacing 39.8 pm . If the reflection from these planes is of the first order, what is the wavelength of the x-rays?
17. The second order Bragg reflection of a crystal is observed at 22° . At what angle will the first order reflection be?

C H A P T E R

7

Superconductivity

The phenomenon of superconductivity was discovered by the Dutch physicist Kamerlingh Onnes while investigating the electrical properties of metals at very low temperatures. Though it was discovered in 1911, understanding of the subject proceeded very slowly. In this chapter, we will consider the basics of superconductivity and some of the phenomena associated with it.

7.1 INTRODUCTION

In 1911, Onnes discovered that for some materials there exists a certain temperature, called **critical temperature** or **transition temperature** T_C , below which the resistivity is zero and the conductivity $\sigma \rightarrow \infty$. This phenomenon is called **superconductivity**. Figure 7.1 shows a plot of the resistance of mercury versus temperature obtained by Onnes. The critical temperature for mercury is 4.2 K. The critical temperature varies from material to material, less than 0.1 K for hafnium to 9.2 K for niobium; below this temperature the electrical resistance of the material is zero. Table 7.1 lists the critical temperatures of some superconducting materials. Many metallic compounds are also superconductors. The superconducting alloy Nb_3Ge , discovered in 1973, has a critical temperature of 23.2 K. This was the highest critical temperature known until 1986.

There is no *d.c.* resistance for a superconductor below T_C . Hence, there can be a current in a ring shaped superconducting material even in the absence of a potential difference. Such currents are called **supercurrents**. In fact, steady currents have been observed to persist for years without loss, in superconducting rings without an electric field.

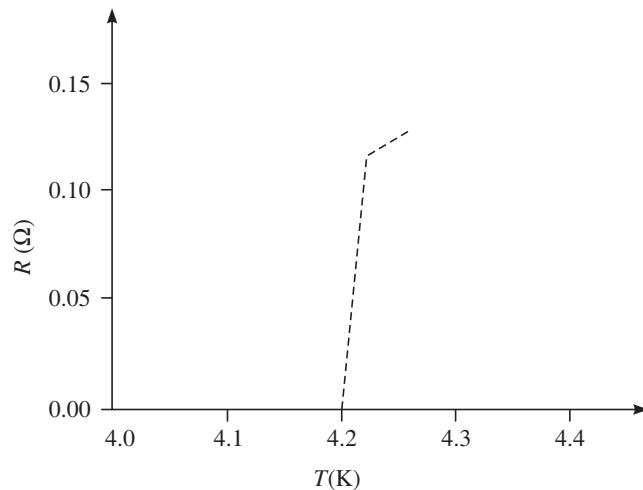


Fig. 7.1 Plot of the resistance of mercury versus temperature, showing sudden decrease of the critical temperature $T = 4.2$ K.

Table 7.1 T_C and B_C values at 0 K for some superconductors

Type-I element*	T_C (K)	B_C (at 0 K) (T)	Type - II compound*	T_C (K)	B_C (at 0 K) (T)
Al	1.175	0.0105	Nb_3Sn	18.1	24.5
Cd	0.517	0.0028	Nb_3Ge	23.2	34.0
Hg	4.154	0.0411	NbN	16.0	15.3
In	3.408	0.0282	V_3Ga	16.5	35.0
Nb	9.25	0.2060	V_3Si	17.1	15.6
Os	0.66	0.0070	Pb MoS	14.4	6.0
Pb	7.196	0.0803	CNb	8.0	1.7
Sn	3.722	0.0305	Al_2CMo_3	9.8	15.6
Tl	2.38	0.0178	Rb_3C_{60}	29.0	?
Zn	0.85	0.0054	$\text{Cs}_2\text{RbC}_{60}$	33.0	?

* For Type - I and Type - II superconductors, See Section 7.4.

7.2 MAGNETIC PROPERTIES

Consider a superconductor at a temperature T below its critical temperature T_C . When a magnetic field B is turned on, the critical temperature is lowered compared to the value when

there is no field. As the magnetic field increases, the critical temperature decreases, as illustrated in Fig. 7.2. If the magnetic field is greater than B_C , called the **critical magnetic field**, superconductivity does not exist at any temperature. In a sense, the above diagram is a phase diagram where the lower left region represents the superconducting state and the upper right region its normal state. Table 7.1 also gives the values of B_C at 0 K.

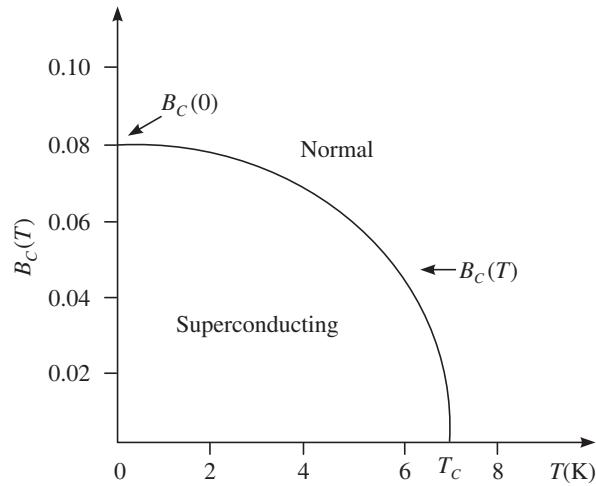


Fig. 7.2 Variation of the critical temperature with magnetic field for lead.

The magnetic field approaches zero as T approaches T_C . The variation of the critical magnetic field at temperature T can be represented by the relation:

$$B_C(T) = B_C(0) \left[1 - \frac{T^2}{T_C^2} \right] \quad (7.1)$$

where $B_C(0)$ is the critical magnetic field at 0 K and T_C is the critical temperature at $B = 0$.

7.3 MEISSNER EFFECT

Another interesting aspect of the magnetic properties of superconductors is the **Meissner effect**. Consider a superconducting material in the presence of a small external magnetic field $B < B_C$ at a temperature greater than the critical temperature. The magnetic field lines will be as illustrated in Fig. 7.3(a). If the material is cooled below the critical temperature T_C , it becomes a superconductor and the magnetic field lines will be as shown in Fig. 7.3(b). In other words, when a superconductor is cooled below the critical temperature in an external magnetic field, the magnetic field lines are expelled or cancelled from the superconductor and the magnetic field inside the superconductor becomes zero.

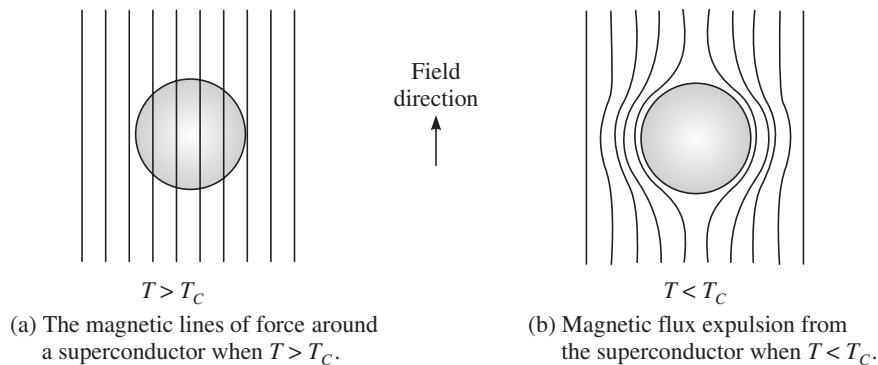


Fig. 7.3 The Meissner effect.

7.4 TYPE-I AND TYPE-II SUPERCONDUCTORS

Depending on the magnetization, superconducting materials can be classified into Type-I and Type-II superconductors. Only certain superconductors, called **type-I or soft superconductors**, exhibit complete Meissner effect. Some of the familiar examples are Al, Cd, Sn, and Zn, (Table 7.1). Figure 7.4(a) shows the plot of the magnetization (M) times μ_0 (permeability of free space) versus the applied magnetic field B for a type-I superconductor. For magnetic fields less than the critical field B_C , the magnetic field induced in the superconductor $\mu_0 M$ is equal and opposite to the external magnetic field. That is, the superconductor is a perfect diamagnet. For type-I superconductors, the values of B_C are very small (Table 7.1). Hence, such materials are not useful as coils of a superconducting magnet.

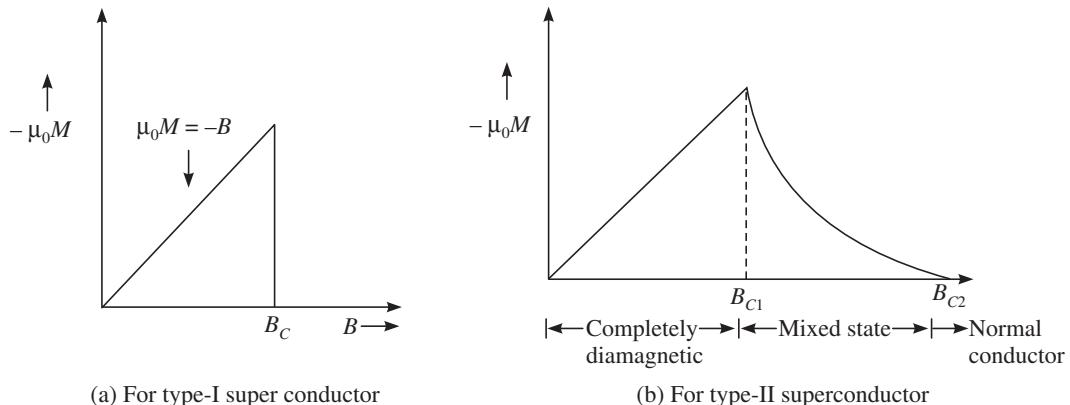


Fig. 7.4 Magnetization versus magnetic field curve.

Type-II or hard superconductors have magnetization curve similar to that given in Fig. 7.4(b). Such materials are usually alloys or metals that have large resistivities in the normal state. Type-II superconductors exhibit two critical magnetic fields, B_{C1} and B_{C2} [Fig. 7.4(b)]. For applied fields less than B_{C1} , total magnetic flux cancellation will be there and the entire

sample will be superconducting as in type-I superconductors. For applied fields greater than B_{C2} , complete penetration of the magnetic field throughout the sample will be there and the resistivity returns to the normal value. For applied fields between B_{C1} and B_{C2} there will be partial penetration of the magnetic fields, the field lines being confined to **flux tubes**, also called **vortices**. In the flux tubes, the material will have normal resistivity. The surrounding material remains field free and superconducting. Each flux tube contains one quantized unit of flux, called a **fluxoid** (Section 7.7).

For certain type-II superconductors, the critical field B_{C2} may be several hundred times larger than the typical values of critical fields for type-I superconductors. For example, Nb_3Ge has a B_{C2} of 34 T . Such materials are used to construct high field superconducting magnets.

Figure 7.5 illustrates the critical magnetic fields B_{C1} and B_{C2} as a function of temperature. Below the B_{C1} curve, the substance exhibits the Meissner effect. Between the two curves is the mixed state with filaments of normal substance penetrating the superconducting state. Above the B_{C2} curve, there is complete magnetic field penetration and the entire sample has normal resistivity.

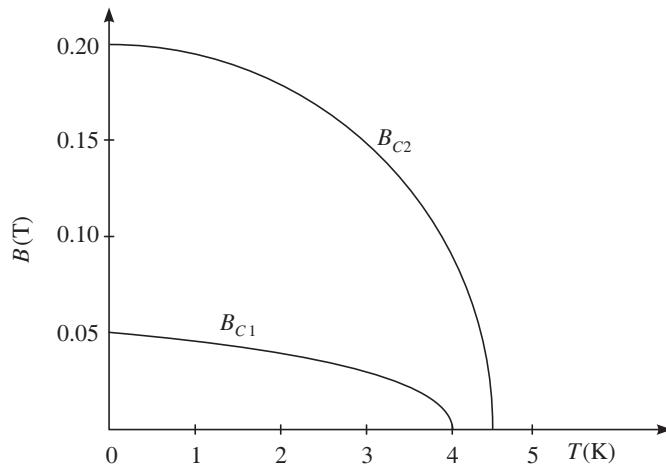


Fig. 7.5 Critical magnetic fields B_{C1} and B_{C2} of a type-II superconductor as a function of temperature.

London penetration depth

In type-I superconductors, magnetic fields are expelled from the interior by the formation of surface currents. In reality, these currents penetrate the surface to a small extent. Within this layer, the magnetic field B decreases exponentially from its external value to zero, according to the expression:

$$B(r) = B_0 \exp\left(-\frac{r}{\lambda}\right) \quad (7.2)$$

where the external field is assumed to be parallel to the surface of the superconductor. B_0 is the magnetic field at the surface, r is the distance from the surface to some interior, and λ is

the **London penetration depth**. The penetration depth varies with temperature according to the equation as follows:

$$\lambda(T) = \lambda_0 \left[1 - \left(\frac{T}{T_c} \right)^4 \right]^{-\frac{1}{2}} \quad (7.3)$$

where λ_0 is the penetration depth at 0 K. From Eq. (7.3), it follows that at $T = T_c$ the substance changes from its superconducting properties and attains its normal state, as the field penetrates the whole specimen.

7.5 THE ISOTOPE EFFECT

It has been observed (1950) that the superconducting critical temperature for different isotopes of superconductors is different. This effect is called the **isotope effect**. Experimental observations revealed that T_c depends on the isotopic mass according to the relation

$$M^\alpha T_c = \text{constant} \quad (7.4)$$

where M is the average isotopic mass and α varies from material to material but is typically about 0.5. The significance of this discovery is that it tells us that the electron-phonon interaction in crystals cannot be ignored.

7.6 BCS THEORY

For quite some time, it was felt that superconductivity is due to a collective behaviour of the conduction electrons. The discovery of isotope effect tells us that the electron-phonon interaction is the one responsible for superconductivity. In 1957, Bardeen, Cooper, and Schrieffer (BCS) published a theory, now known as **BCS theory**, which explained most of the phenomena relating to superconductivity.

7.6.1 Cooper Pairs

According to BCS theory, at low temperatures the conduction electrons in a superconductor are coupled in pairs. The coupling comes in the following way:

An electron moving through the lattice of positive ions attracts the positive ions nearby, displacing them slightly, resulting in a region of increased positive charge density (Fig. 7.6). This region of increased charge density propagates through the material as a vibrational wave in the lattice, called a **phonon**.

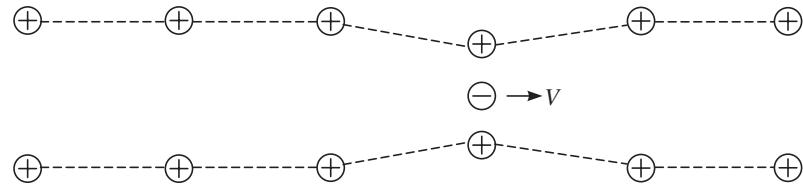


Fig. 7.6 An electron travelling through the lattice of positive ions generating a wave of increased charge density, shown by the dotted lines in two dimensions.

A second electron that encounters the wave of increased positive charge concentration is attracted towards it by the Coulomb interaction. Thus, the phonon acts as the mediator of the interaction between the two electrons. Since both the electrons experience an attractive force toward the region of increased positive charge density, the interaction between the two is an attractive one. At low temperature ($T < T_C$), the attraction between the two electrons can exceed the Coulomb repulsion between them. This pair of electrons is called a **Cooper pair**.

7.6.2 Superconducting Energy Gap

The electrons in a Cooper pair have opposite spins and equal and opposite linear momenta. Thus, they form a system with zero spin and zero linear momentum. Each Cooper pair may be considered as a single particle with zero spin. Collection of such Cooper pairs form a system of bosons. Hence, any number of Cooper pairs may be in the same quantum state with the same energy. In the superconducting state, the Cooper pairs are correlated, so that they all act together. Thus, in the ground state of a superconductor (at $T = 0$), all the electrons are in Cooper pairs and all the Cooper pairs are in the same state. The electrons in a superconducting state can absorb or emit energy when the binding of the Cooper pairs is broken. The energy needed to breakup a Cooper pair is similar to that needed to break up a molecule into its constituent atoms. This energy is called the **superconducting energy gap** E_g . According to BCS theory this energy at absolute zero is predicted to be

$$E_g = 3.5 k T_C \quad (7.5)$$

where k is the Boltzmann constant. As the temperature is increased from $T = 0$ K, the energy gap reduces. At $T = T_C$ the energy gap is zero.

The energy gap for typical superconductors ($\sim 10^{-4}$ eV) are much smaller than the energy gap for typical semiconductors (~ 1 eV). This prediction of the energy gap was verified experimentally by studying the absorption of electromagnetic radiation by superconductors.

7.7 MAGNETIC FLUX QUANTIZATION

Consider a ring made up of a superconductor in its normal state in a magnetic field illustrated in Fig. 7.7(a). Cool the ring to a temperature below the critical temperature T_C . The magnetic field in the material of the ring will be expelled. Some of the expelled magnetic field lines will be trapped inside the ring as in Fig. 7.7(b). A prediction of the BCS theory is that the magnetic flux inside the ring is quantized, taking on the values:

$$\Phi = n \frac{h}{2e} \quad (7.6)$$

where n is an integer and $2e$ refers to the charge on the Cooper pair that forms the basic unit. The magnitude of the flux quantum is given by

$$\Phi_0 = \frac{h}{2e} = 2.0678 \times 10^{-15} \text{ Tm}^2 \quad (7.7)$$

This is referred as the **quantum fuxoid**. Quantization of magnetic flux has been confirmed experimentally.

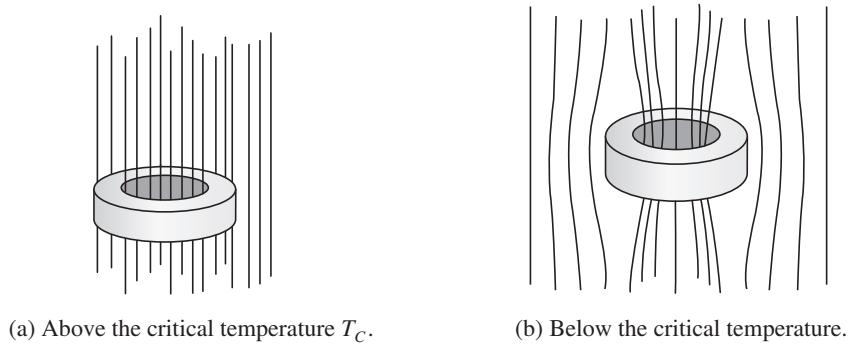


Fig. 7.7 A ring of superconducting material in a magnetic field.

7.8 JOSEPHSON JUNCTION

In barrier penetration (to be discussed in Section 9.22) a single particle tunnels through a barrier. Consider two superconductors separated by a thin layer ($\sim 10^{-9}$ m) of insulating material. In 1962, Josephson proposed that Cooper pairs could tunnel from one superconductor to the other with no resistance, giving rise to a direct current with no voltage across the junction. Such a superconductor-insulator-superconductor junction is called a **Josephson junction** (Fig. 7.8).

The d.c. Josephson effect

The current across the junction is given by

$$I = I_{\max} \sin(\phi_2 - \phi_1) \quad (7.8)$$

where I_{\max} is the maximum current which is dependent on the thickness of the barrier, ϕ_1 is the phase of the wave function on one side of the insulating layer, and ϕ_2 is the phase of the corresponding wave function on the other side. This result is known as the **d.c. Josephson effect**.

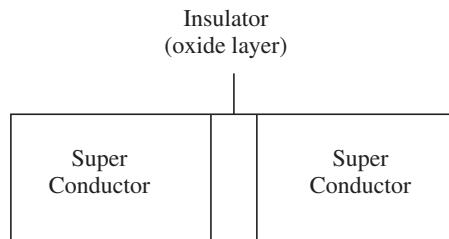


Fig. 7.8 Josephson junction.

The a.c. Josephson effect

If a *d.c.* voltage V is applied across a Josephson junction, there would be an alternating current, having frequency v given by

$$v = \frac{2 eV}{h} \quad (7.9)$$

This result is known as the **a.c. Josephson effect**, and the current is called the **Josephson current**. As frequency can be measured very accurately, a.c. Josephson effect is often used as a voltage standard. The accuracy of the voltage standard with Josephson junction is approximately one part in 10^{10} . The inverse effect also happens, in which the application of an alternating voltage across a Josephson junction results in a *d.c.* current. This has also been verified experimentally.

7.9 HIGH TEMPERATURE SUPERCONDUCTIVITY

Superconductors can be grouped into low T_C and high T_C based on their transition temperature. Substances having T_C around or below 24 K are considered as low T_C super-conductors and those above are **high T_C superconductors**. The low T_C superconductors range from ordinary metals and alloys such as Nb($T_C = 9.25$ K) and Nb₃Ge($T_C = 23.2$ K) to oxides like BaPb_{1-x}Bi_xO₃($T_C = 13$ K). A class of compounds called **A-15**, Nb₃Ge, Nb₃Sn($T_C = 18.1$ K) and Nb₃Ga($T_C = 20.3$ K) and others, are the very familiar ones with T_C in and around 20 K. The BCS theory could explain the behaviour of low T_C superconductors.

A forward step was the observation of a T_C of 40 K in La_{1-x}Sr_xCuO with $x = 0.15$ by Bednorz and Muller (1986), for which they were awarded the Nobel Prize in 1987. Changing lanthanum by yttrium and strontium by barium a T_C of 93 K in YBa₂Cu₃O₇ was reported in 1987. This compound is referred to as **1-2-3** following the ratio of the three metallic elements. This was a fantastic advancement, as the observed T_C surpassed 77 K, the boiling point of liquid nitrogen. Another interesting aspect about the type-II copper oxide superconductors is about their extremely high upper critical fields. For Y Ba₂Cu₃O₇ the upper critical field $B_{c2} \approx 100$ T at 77 K and about 300 T at 0 K.

A similar compound using bismuth and aluminum Bi Al_{1-y}CaSrCoO_{7-d} with $d < 0.45$ and $0 < y < 0.3$ raised the value of T_C to 114 K. The thallium based compound Tl₂ Ba₂Ca_{n-1} Cu_nO_{2n+4} is a very interesting since it gives different T_C values for different n ; T_C of 80 K for $n = 1$, $T_C = 110$ K for $n = 2$ and $T_C = 125$ K for $n = 3$. There is a possibility of reaching very high T_C values if we can make compounds with $n > 3$. In 1993, a higher T_C (113 K) was achieved in the mercury based copper oxide HgBa₂Ca₂Cu₃O_{1+x} with x being a small positive number. At extremely high pressures (about 300000 atmospheres) these mercury based superconductors reach transition temperature of 164 K. Most of these compounds fall in the category of ceramics.

Another class of superconductors is based on the organic molecule C₆₀. Pure C₆₀ is not a superconducting element. When C₆₀ is doped with the right amount of potassium, it forms the compound K₃C₆₀ with a T_C of 18 K. Combination of C₆₀ with thallium and rubidium increases the T_C to 42.5 K. The dream of the scientists is to achieve a T_C of about 300 K, so that they can eliminate the need for cryogenic fluids.

7.10 APPLICATIONS

The properties of zero resistance and the Meissner effect make superconductors ideal for many applications. The cost of cryogenic systems and fluids has also to be taken into account when we consider the applications. We discuss here some of the important applications.

7.10.1 SQUID

The Josephson junctions are used in devices known as Superconducting Quantum Interference Devices (SQUID). The SQUID uses a pair of Josephson junctions in a current loop. The current is extremely sensitive to the magnetic flux applied to the loop. SQUIDS are used for measuring very small amounts of magnetic flux. They can be used to measure quantum fluxoid to within one part in 10^6 . As ordinary magnetometers, SQUIDS are capable of measuring magnetic fluctuations of the order of 10^{-13} T. The biomedical applications of SQUIDS include the imaging of soft tissues such as the brain. They have also been used to detect the presence of bacteria, search for explosives, examine materials for defects, and search for oil.

7.10.2 Generation and Transmission of Electricity

Substantial amount of energy saving is possible if superconductors are used in electrical generators and motors. Most of the savings come from the replacement of the heavy iron cores by lighter superconducting magnets. The other part comes from the cutting of resistive losses. Superconducting transmission lines would save significant amounts of energy. Superconducting transmission requires no expensive transformers. For normal transmission one has to make the voltage as large as possible, which is not needed if $R = 0$. The large current densities possible in superconducting wires allow the reduction of the number and size of transmission lines. Superconducting energy storage rings would be of great use to store extra energy generated during low usage hours in power plants.

7.10.3 Superconducting Magnets

A conventional electromagnet is made by passing current through a metal wire wound in the shape of a solenoid. A maximum magnetic field of a few millitesla can be generated by this method. By placing an iron core within the solenoid, the magnetic field can be enhanced to about 2 to 3 T. The iron core is extremely heavy and cumbersome. Since we can achieve current densities of 10^7 A/cm² in superconducting wires, it is possible to produce very large magnetic fields. An additional advantage is that no iron core is needed in this case. The superconducting material must have a critical field higher than the maximum field we want to generate. With type-II superconductors, it is possible to achieve upper critical fields of 50T (Section 7.4). Superconducting magnets are used for large particle accelerators. Large magnets are used to confine plasma in fusion research. One of the most important medical applications of large superconducting magnets is in Magnetic Resonance Imaging (MRI).

7.10.4 Other Applications

Josephson junctions could be used in integrated circuits. Superconducting computers are not cost-effective. They may have to be re-evaluated when room temperature superconductors

become available. The Josephson junction is used to measure and maintain voltage standards. In Japan, a magnetic levitation (maglev) train achieved a speed of 550 km/h. It is designed on the principle of Meissner effect to levitate the transport system. The idea is to make a more comfortable travel at higher speeds.

SOLVED EXAMPLES

Example 7.1 A superconducting sample has a critical temperature of 3.722 K in zero magnetic field and a critical field of 0.0305 T at 0 K. Evaluate the critical field at 2 K.

$$\begin{aligned}\textbf{Solution:} \quad & \text{We have } B_C(T) = B_C(0) \left[1 - \frac{T^2}{T_C^2} \right] \\ & B_C(2\text{ K}) = (0.0305 \text{ T}) \left[1 - \left(\frac{2\text{ K}}{3.722\text{ K}} \right)^2 \right] \\ & = 0.0217 \text{ T}\end{aligned}$$

Example 7.2 A *d.c.* voltage of 1 μ V is applied across a Josephson junction. Calculate the frequency of the Josephson current generated.

Solution: From Eq. (7.9), the required frequency

$$\begin{aligned}v &= \frac{2eV}{h} = \frac{2(1.6 \times 10^{-19} \text{ C})(10^{-6} \text{ V})}{6.626 \times 10^{-34} \text{ Js}} \\ &= 482.9 \text{ MHz}\end{aligned}$$

Example 7.3 Calculate the superconducting energy gap at $T = 0$ K predicted by the BCS theory for cadmium. T_C for cadmium is 0.517 K.

Solution: Energy gap $E_g = 3.5 \text{ K } T_C$

$$\begin{aligned}E_g &= \frac{3.5(1.38 \times 10^{-23} \text{ J/K})(0.517 \text{ K})}{1.6 \times 10^{-19} \frac{\text{J}}{\text{eV}}} \\ &= 1.56 \times 10^{-4} \text{ eV}\end{aligned}$$

Note: Compare this result with the measured value of 1.5×10^{-4} eV.

Example 7.4 The superconducting energy gap for a material is 1.5×10^{-4} eV. Compute the wavelength of a photon whose energy is just sufficient to break up a Cooper pair in it.

Solution: Let v be the frequency of the photon. Then,

$$E_g = h\nu = \frac{hc}{\lambda} \quad \text{or} \quad \lambda = \frac{hc}{E_g}$$

$$\begin{aligned}\lambda &= \frac{(6.626 \times 10^{-34} \text{ Js})(3 \times 10^8 \text{ m/s})}{(1.5 \times 10^{-4} \text{ eV})(1.6 \times 10^{-19} \text{ J/eV})} \\ &= 8.28 \times 10^{-3} \text{ m}\end{aligned}$$

Example 7.5 Calculate the London penetration depth for lead at 5.2 K if the penetration depth at 0 K is 37 nm. The critical temperature of lead is 7.193 K.

Solution: London penetration depth is given as

$$\begin{aligned}\lambda(T) &= \lambda_0 \left[1 - \left(\frac{T}{T_c} \right)^4 \right]^{-\frac{1}{2}} \\ \lambda(5.2 \text{ K}) &= 37 \text{ nm} \left[1 - \left(\frac{5.2 \text{ K}}{7.193 \text{ K}} \right)^4 \right]^{-\frac{1}{2}} = 43.4 \text{ nm}\end{aligned}$$

Example 7.6 An isotope of mercury with a mass of 199 amu has a transition temperature of 4.185 K. If the T_C of another isotope is 4.153 K, what is its atomic mass. Take the value of α to be 0.5.

Solution: We have the relation

$$M^\alpha T_C = \text{Constant}$$

$$\frac{M_2}{M_1} = \left(\frac{T_{C1}}{T_{C2}} \right)^\frac{1}{\alpha}$$

$$\begin{aligned}M_2 &= (199 \text{ amu}) \left(\frac{4.185}{4.153} \right)^2 \\ &= 202.08 \text{ amu}\end{aligned}$$

REVIEW QUESTIONS

- 7.1 Explain what is superconductivity.
- 7.2 Explain critical temperature and critical magnetic field with respect to a superconducting substance. State the relation connecting the two.

- 7.3** What is Meissner effect?
- 7.4** What are type-I and type-II superconductors? Give examples.
- 7.5** Type-I superconductors are not suitable as coils of superconducting magnet whereas type-II are. Why?
- 7.6** Explain what you understand by vortices in type-II superconductors.
- 7.7** What is isotope effect in superconductivity? Explain its importance.
- 7.8** What are Cooper pairs?
- 7.9** What is Josephson tunnelling? Mention some of the applications of Josephson junction.
- 7.10** Write a note on high temperature superconductors.
- 7.11** Explain one or two applications of superconductivity.
- 7.12** What is *a.c.* Josephson effect? How is it used as a voltage standard?
- 7.13** Explain briefly magnetic flux quantization.
- 7.14** Explain Meissner effect and type-I and type-II superconductors. Write a note on high temperature superconductors.
- 7.15** How is superconductivity explained on the basis of the BCS theory? Explain some of the applications of superconductors.
- 7.16** Discuss the magnetic properties of superconductors.

PROBLEMS

1. A superconducting sample has a critical temperature of 4.1 K in zero magnetic field and critical magnetic field of $0.0505 T$ at 0 K. Find the critical field at 2 K.
2. Estimate the critical field for lead at 3.8 K from the following data.
 $B_C(0) = 5.6 \times 10^4$ A/m, and $T_C = 7.193$ K.
3. The critical temperature of a superconducting sample is 1.19 K with atomic mass 26.9. Estimate the critical temperature when the isotope mass change to 32.1.
4. Compute (i) the superconducting energy gap for indium and (ii) the wavelength of a photon that could just break up a Cooper pair in indium. The transition temperature of indium is 3.408 K.
5. The London penetration depth of lead (Pb) at 0 K is 390 Å. Calculate the penetration depth at 2 K. T_C of lead is 7.2 K.
6. Calculate the wavelength of the photon, which will be required to break a Cooper pair in the superconductor Zr whose $T_C = 0.56$ K.

C H A P T E R

8

Special Theory of Relativity

The theories developed during three centuries from 1600 had been very successful in explaining most of the phenomena in physical science. The Newtonian mechanics explained the dynamics of objects on earth and in the heavens. It also successfully explained wave motion and the behaviour of fluids. The kinetic theory of matter showed the connection between mechanics and heat. Maxwell's electromagnetic theory unified the branches of optics, electricity, and magnetism into a single large field called **electrodynamics**. Towards the close of the nineteenth century certain new discoveries (x-rays, radioactivity, and electron) and experimental observations (blackbody radiation curves, photoelectric effect, optical spectra, etc.) were discovered. These revealed many phenomena the existing theories failed to explain. However, the physicists were firm and confident of explaining the old theories on the basis of present theories. However, with the formulation of two revolutionary new theories, the quantum theory and the theory of relativity, they were convinced about the inability of classical physics to explain all the physical phenomena. In this chapter we discuss the special theory of relativity, which was proposed by Albert Einstein, in 1905.

8.1 INERTIAL AND NONINERTIAL REFERENCE FRAMES

The basic concepts for the study of motion are space and time, both of which are assumed to be continuous. To describe the motion of a body, one has to specify its position in space as a function of time. To do this, a co-ordinate system is used as a frame of reference. The reference frames in which Newton's law of inertia holds good are called **inertial reference frames**. The remaining laws are also valid in inertial reference frames only. The acceleration of an inertial reference frame is zero, and therefore, it moves with a constant velocity. Any reference frame that moves with constant velocity relative to an inertial frame is also an inertial frame of

reference. For simple applications in the laboratory, reference frames fixed on the earth are inertial frames. For astronomical applications, the terrestrial frame cannot be regarded as an inertial frame. Reference frames where the law of inertia does not hold are called **noninertial reference frames**.

8.2 GALILEAN TRANSFORMATION

Let us consider two inertial systems S and S' with co-ordinate axes xyz and $x'y'z'$ attached to them. Let S' move with respect to S with a uniform velocity v along the xx' axes as shown in Fig. 8.1. The origins of the two systems coincide when $t = t' = 0$. Let an event take place at P whose co-ordinates with respect to S be (x, y, z, t) and with respect to S' be (x', y', z', t') . From Fig. 8.1, it is obvious that these co-ordinates are related by

$$x' = x - vt \quad (8.1)$$

$$y' = y \quad (8.1a)$$

$$z' = z \quad (8.1b)$$

$$t' = t \quad (8.1c)$$

These are called the **Galilean transformation equations** or **Newtonian transformation equations**.

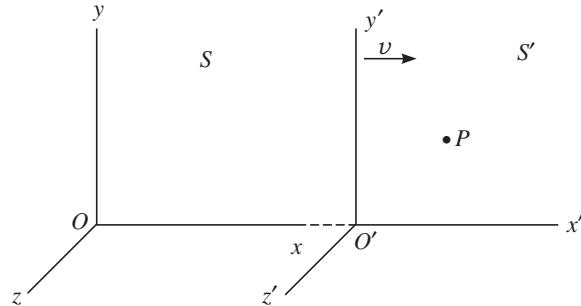


Fig. 8.1 The inertial systems S and S' with co-ordinate axes xyz and $x'y'z'$.

The **Galilean velocity transformation** is obtained by differentiating Eqs. (8.1 – 8.1b) with respect to time and using the result $\left(\frac{d}{dt}\right) = \left(\frac{d}{dt'}\right)$:

$$\begin{aligned} \frac{dx'}{dt'} &= \frac{dx}{dt} - v & u'_x &= u_x - v \\ \frac{dy'}{dt'} &= \frac{dy}{dt} & \text{or} & u'_y = u_y \end{aligned} \quad (8.2)$$

$$\frac{dz'}{dt'} = \frac{dz}{dt} \quad u'_z = u_z$$

Since v is constant, from Eq. (8.2), it is evident that acceleration $a'_x = a_x$, $a'_y = a_y$, $a'_z = a_z$. That is, acceleration is invariant with respect to Galilean transformation. In Newtonian formulation, mass is absolute, and therefore, force (which is mass \times acceleration) is absolute. This implies that Newton's second law is invariant with respect to Galilean transformation.

Newtonian relativity assumes that space and time are absolute quantities. Their measurement does not change from one inertial frame to another. The mass of an object and force are unchanged by a change in the inertial frame. But the position of an object and its velocity are different in different inertial frames.

8.3 ELECTROMAGNETISM AND GALILEAN TRANSFORMATION

Maxwell's equations predicted the existence of electromagnetic waves propagating through space with a speed of 3×10^8 m/s. Then, a spherical electromagnetic wave propagating with a constant speed c in the reference frame S is given by

$$x^2 + y^2 + z^2 - c^2 t^2 = 0 \quad (8.3)$$

For Eq. (8.3) to be invariant, its form in the system S' should be

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = 0 \quad (8.4)$$

Substituting the values of x , y , z , t in Eq. (8.3) in terms of x' , y' , z' , t' , we get

$$(x' + vt)^2 + y'^2 + z'^2 - c^2 t'^2 = 0 \quad (8.5)$$

which is not the same as Eq. (8.4). Hence, the Galilean transformation equations do not hold good in the case of electromagnetism.

8.3.1 Ether Hypothesis

The nineteenth century physicists used to view various phenomena in terms of the laws of mechanics. The mechanical wave phenomena require a medium to support the wave. Therefore, physicists assumed that light and other electromagnetic waves must travel in some medium. They called this transparent medium the **ether**, and assumed that it permeates all space. They had to assign very strange properties to ether. It had to be transparent and massless so that electromagnetic waves could travel through vacuum. On the contrary, it had to be very hard to support the transverse vibrations of the wave motion. The ether hypothesis led to the following two alternatives:

- (i) The **stationary ether hypothesis**, wherein the ether is at rest with respect to the bodies moving through it. The reference frame wherein the ether could be considered at rest is called the **ether frame** or the **rest frame** or **absolute frame**. In this frame the velocity of light is always c .

- (ii) The **ether drag hypothesis**, wherein ether is dragged along with the bodies that move through it.

A number of experiments were designed to check the ether hypothesis. Of these, the most direct one is the one performed by Michelson and Morley in the 1880's.

8.4 MICHELSON-MORLEY EXPERIMENT

The purpose of the Michelson-Morley experiment was to confirm the existence of an absolute frame of reference (stationary ether). If the ether is at rest, when the earth moves through it there must be a relative velocity of the earth with respect to the ether. What they did was to measure the difference in the speed of light in different directions.

8.4.1 The Interferometer

The experimental set up used is the same as the one discussed in Section 3.10, which is again shown in Fig. 8.2. The light from a source S is split into two beams by a half silvered plate P . One beam travels to mirror M_1 and the other to mirror M_2 . The beams are reflected by M_1 and M_2 and are recombined again after passing through P . Beam 2 goes through the plate P three times, whereas beam 1 goes through P only once. Hence, to make the optical paths of the two beams equal, a compensating plate P' is placed in the path of Beam 1. Beams 1 and 2 arrive at the telescope T and produce interference fringes. If the optical path lengths of the beams are exactly equal, constructive interference occurs, leading to a bright fringe. If one mirror is moved a distance $\lambda/4$, which corresponds to a path difference of $\lambda/2$ between the beams, destructive interference occurs giving rise to a dark fringe. Thus, by moving one of the mirrors, the fringe system can be made to move past a crosswire which serves as the reference mark. Let us assume that the earth is moving to the right with a velocity v with respect to the stationary ether (Fig. 8.2). Michelson arranged the interferometer in such a way that PM_1 is parallel to the direction of the velocity v . To reduce mechanical vibrations, the interferometer was mounted on a large stone that floated in a tank of mercury.

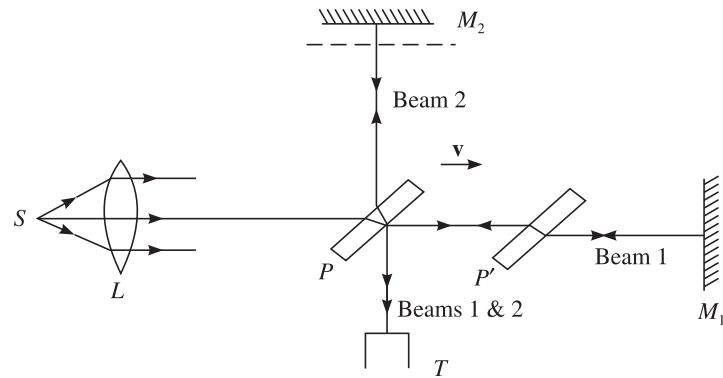


Fig. 8.2 Schematic representation of the Michelson—Morley experiment.

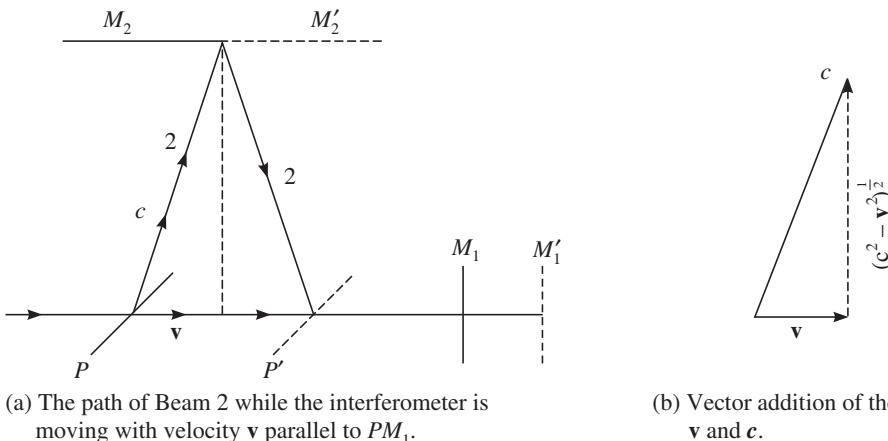
8.4.2 The Experiment

To start with, mirrors M_1 and M_2 are set such that $PM_1 = PM_2 = d$. If the apparatus is stationary in ether the two waves take the same time to return to the telescope, and hence, meet in the same phase. But the apparatus is moving with the same velocity \mathbf{v} to the right. Therefore, the time required by the two waves for the to and fro journey through the same distance will not be equal. First, we consider beam 1 which travels parallel to the velocity \mathbf{v} . The transmitted wave travels towards M_1 with relative velocity $c - \mathbf{v}$. Hence, the time required by this wave for its round trip is:

$$t_1 = \frac{d}{c - \mathbf{v}} + \frac{d}{c + \mathbf{v}} = \frac{2cd}{c^2 - \mathbf{v}^2} = \frac{2d}{c \left(1 - \frac{\mathbf{v}^2}{c^2}\right)} \quad (8.6)$$

The path of beam 2 when the interferometer is moving with velocity \mathbf{v} parallel to beam 1 is illustrated in Fig. 8.3(a). By vector addition, the velocity component perpendicular to the direction of motion of the interferometer is $(c^2 - \mathbf{v}^2)^{\frac{1}{2}}$ [Fig. 8.3(b)]. The time taken by beam 2 to travel from P to M_2 and back is:

$$t_2 = \frac{2d}{(c^2 - \mathbf{v}^2)^{\frac{1}{2}}} = \frac{2d}{c \left(1 - \frac{\mathbf{v}^2}{c^2}\right)^{\frac{1}{2}}} \quad (8.7)$$



(a) The path of Beam 2 while the interferometer is moving with velocity \mathbf{v} parallel to PM_1 . (b) Vector addition of the velocities \mathbf{v} and \mathbf{c} .

Fig. 8.3

The difference between the time taken by the two beams is given by

$$\Delta t = t_1 - t_2 = \frac{2d}{c} \left[\left(1 - \frac{\mathbf{v}^2}{c^2}\right)^{-1} - \left(1 - \frac{\mathbf{v}^2}{c^2}\right)^{-\frac{1}{2}} \right]$$

Since $\frac{\mathbf{v}}{c} \ll 1$, using binomial expansion, we get

$$\Delta t = \frac{2d}{c} \left[\left(1 + \frac{\mathbf{v}^2}{c^2} + \dots \right) - \left(1 + \frac{\mathbf{v}^2}{2c^2} + \dots \right) \right] \approx \frac{d\mathbf{v}^2}{c^3} \quad (8.8)$$

If $\mathbf{v} = 0$, then $\Delta t = 0$ and the two beams take the same time.

In their experiment, Michelson and Morley rotated the interferometer through an angle of 90° . In the rotated position, beam 2 will be parallel to the velocity \mathbf{v} and beam 1 will be perpendicular to it. This led to a total difference in time $(\Delta t) + (\Delta t)$, which was equivalent to

a path difference of $2(\Delta t)c = \frac{2d\mathbf{v}^2}{c^2}$. Therefore, there would be a shift in the fringe system across the crosswire in the telescope. The number of fringes shifted is given as

$$\Delta N = \frac{\text{Path difference}}{\text{Wave length}} = \frac{\frac{2d\mathbf{v}^2}{c^2}}{\lambda} = \frac{2d\mathbf{v}^2}{c^2\lambda} \quad (8.9)$$

8.4.3 Results

Michelson and Morley reflected beams 1 and 2 back and forth 8 times, so that the path length becomes 10 m. With $\lambda = 5000 \text{ \AA}$, Eq. (8.9) leads to a fringe shift given by

$$\Delta N = \frac{2(10 \text{ m}) (3 \times 10^4 \text{ ms}^{-1})^2}{(3 \times 10^8 \text{ ms}^{-1})^2 (5 \times 10^{-7} \text{ m})} = 0.4 \quad (8.10)$$

Detection of this fringe width would be possible since their apparatus was capable of observing a fringe shift of as small as 0.01 fringe. But they could not observe any significant fringe shift. They repeated the experiment at different places at different times of the day in different seasons. Still they could not observe a significant fringe shift, indicating the absence of ether, and found that the speed of light in the interferometer is the same for the two perpendicular paths. This null result remained as one of the great puzzles of physics towards the end of 19th century.

The null result can be explained if the Galilean transformation is abandoned and the velocity of light is assumed to be the same in all inertial frames. Then, for beam 1, $t_1 = 2 \frac{d}{c}$ and for beam 2, $t_2 = \frac{2d}{c}$ which leads to $\Delta t = t_1 - t_2 = 0$. It is also evident from Eq. (8.9) that ΔN approaches zero, if $\mathbf{v} \ll c$. Hence, we may consider that the Galilean transformation is valid when $\mathbf{v} \ll c$. In other words, Galilean transformation is valid for mechanics, but not for electromagnetism since the velocity of electromagnetic waves is equal to c .

The results of Michelson and Morley were a real challenge and a number of explanations were put forth over a period of years. The radical new theory proposed by Einstein in 1905 explained various experimental results satisfactorily, and changed our ideas about space and time.

8.5 THE POSTULATES OF SPECIAL THEORY OF RELATIVITY

Einstein in his theory dropped the concept of ether and the accompanying assumption of an absolute frame of reference at rest. Also, he revised the classical ideas regarding space and time by asserting that absolute motion is meaningless. Einstein's ideas are embodied in two postulates which are as follows:

Postulate 1: The principle of equivalence

"The laws of physics have the same form in all inertial reference frames."

Postulate 2: Constancy of the speed of light

"The speed of light in free space (vacuum) is always a constant c and is independent of the speed of the source or observer or the relative motion of the inertial systems."

These two postulates form the foundation of Einstein's special theory of relativity. It is referred to as special to distinguish it from his later theory, the general theory of relativity which deals with noninertial frames.

8.6 LORENTZ TRANSFORMATION

The transformation equations for inertial frames of reference moving with uniform relative velocity were derived by Einstein. However, they are called Lorentz transformations since Lorentz derived (1890) the same relations in electromagnetism. Let us consider two reference frames S and S' moving with uniform relative motion as described in Section 8.2. Let two observers O and O' observe any event P from systems S and S' , respectively. Let the event P be produced at $t = 0$ when the origins of the two frames coincide. For the observer at O , let the co-ordinates of the event at a particular instant be (x, y, z, t) . The same event is described by the co-ordinates (x', y', z', t') for the observer O' on the system S' . Using the two postulates and assuming that the transformation must reduce to Galilean transformation at low speed, Einstein derived the following transformation equations between (x, y, z, t) and (x', y', z', t') :

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (8.11)$$

$$y' = y \quad (8.11a)$$

$$z' = z \quad (8.11b)$$

$$t' = \frac{t - \left(\frac{vx}{c^2}\right)}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (8.11c)$$

Equations (8.11–8.11c) are called the **Lorentz transformation** equations which Einstein derived in electromagnetism. Here, it is done on a more dynamical basis.

The inverse transformation can be obtained by interchanging the primed and unprimed quantities, and reversing the sign of the relative velocity since S and S' differ only in the sign of relative velocity. They are

$$x = \frac{x' + vt}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (8.12)$$

$$y = y' \quad (8.12a)$$

$$z = z' \quad (8.12b)$$

$$t = \frac{t' + \left(\frac{vx}{c^2} \right)}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (8.12c)$$

With the usual abbreviations

$$\beta = \frac{v}{c}; \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \beta^2}} \quad (8.13)$$

the Lorentz transformation and the inverse transformation simplify as

$$x' = \gamma(x - \beta ct) \quad x = \gamma(x' + \beta ct') \quad (8.14)$$

$$y' = y \quad y = y' \quad (8.14a)$$

$$z' = z \quad \text{or} \quad z = z' \quad (8.14b)$$

$$t' = \gamma \left(t - \frac{\beta}{c} x \right) \quad t = \gamma \left(t' + \frac{\beta}{c} x' \right) \quad (8.14c)$$

In the low velocity limit, $\beta \ll 1$; it follows that the Lorentz transformation reduces to the Galilean transformation. If $v > c$, the quantity $\sqrt{1 - \beta^2}$ becomes imaginary. The space and time co-ordinates would then become imaginary, which is physically unacceptable. Hence, in vacuum nothing can move with a velocity greater than the velocity of light. It may be noted that the space transformation, Eq. (8.11a) involves time and the time transformation Eq. (8.11c) involves the space co-ordinate. Hence, the transformation is sometimes referred to as **space-time transformation**.

8.7 SIMULTANEITY

An important consequence of Lorentz transformation is that simultaneity is relative. Let us consider two events occurring at two different points x_1 and x_2 at times t_1 and t_2 in the inertial system S . Let t'_1 and t'_2 be the times at which the two events are observed to occur with respect to S' . Then from Lorentz transformation, we get

$$t'_1 = \gamma \left(t_1 - \frac{\mathbf{v}x_1}{c^2} \right), \quad t'_2 = \gamma \left(t_2 - \frac{\mathbf{v}x_2}{c^2} \right)$$

$$t'_2 - t'_1 = \gamma (t_2 - t_1) + \frac{\gamma \mathbf{v}}{c^2} (x_1 - x_2) \quad (8.15)$$

If the two events are occurring at the same instant in S , $t_2 - t_1 = 0$, and

$$t'_2 - t'_1 = \frac{\gamma \mathbf{v}}{c^2} (x_1 - x_2) \neq 0 \quad (8.16)$$

That is, two events that are simultaneous in one reference frame are not simultaneous in another frame of reference moving relative to the first, unless the two events occur at the same point in space. It implies that clocks that appear to be synchronized in one frame of reference will not necessarily be synchronized in another frame of reference in relative motion.

8.8 TIME DILATION

Let us consider two successive events occurring at the same point x' in the inertial frame S' . Let t'_1 and t'_2 be the times recorded by the observer in frame S' . Then the time interval measured is $t'_2 - t'_1$. For the observer in S' , the rest frame is S' itself. The time interval between events in the rest frame, is called the **proper time** $\Delta\tau$. Hence,

$$\Delta\tau = t'_2 - t'_1 \quad (8.17)$$

However, an observer in frame S measures this instant as:

$$t_1 = \gamma \left[t'_1 + \left(\frac{\mathbf{v}x'}{c^2} \right) \right]; \quad t_2 = \gamma \left[t'_2 + \left(\frac{\mathbf{v}x'}{c^2} \right) \right]$$

The time interval according to the observer in frame S is then $\Delta t = t_2 - t_1$, which is given by

$$\Delta t = \gamma (t'_2 - t'_1) \quad \text{or} \quad \Delta t = \gamma \Delta\tau$$

$$\Delta t = \frac{\Delta\tau}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} = \frac{\Delta\tau}{\sqrt{1 - \beta^2}} \quad (8.18)$$

Since $\gamma > 1$, it follows from Eq. (8.18) that the proper time interval is a minimum. The effect is known as **time dilation** and is equivalent to the slowing down of moving clocks. Hence, growth, aging, pulse rate, heart beats etc., are slowed down in a fast moving frame. If the velocity of the moving frame $v = c$, $\Delta t \rightarrow \infty$ and the process of aging will stop altogether.

The time dilation effect has been verified experimentally by observation on elementary particles and by atomic clocks accurate to nanoseconds that are carried aboard jet planes.

8.9 LENGTH CONTRACTION

Let S and S' be inertial systems moving with relative velocity \mathbf{v} along the xx' axes. Consider a rod at rest in the inertial system S' lying parallel to the x' axis. Although the system S' is moving with a relative velocity, to the observer in S' the rod is at rest. The length in an inertial frame in which the rod is at rest is called its **proper length**. The length of the rod $L_0 = x'_2 - x'_1$, where x'_1 and x'_2 are the co-ordinates of its two ends measured at the same instant of time. To an observer in S , the length of the rod $L = x_2 - x_1$, where x_1 and x_2 are the co-ordinates of its two ends measured at the same time $t_2 = t_1 = t$. Using Lorentz transformation, we get

$$x'_2 = \gamma(x_2 - vt_2); \quad x'_1 = \gamma(x_1 - vt_1)$$

Use of these expressions in $L_0 = x'_2 - x'_1$ gives,

$$\begin{aligned} L_0 &= \gamma(x_2 - x_1) = \gamma L \\ L &= L_0 \sqrt{1 - \beta^2} \end{aligned} \tag{8.19}$$

Since $\sqrt{1 - \beta^2}$ is always less than unity, the length $L < L_0$. That is, to an observer in S the rod looks as though it is contracted parallel to the direction of motion.

The effect is reciprocal. If a rod has a length L_0 in S , to an observer in S' which is in relative motion, it will appear to be of length $L_0 \sqrt{1 - \beta^2}$. The phenomenon of length contraction is referred to as **Lorentz-Fitzgerald contraction**. Thus, space which is reduced to the measurement of length in physics and the geometrical shapes of objects cannot be absolute, but only relative.

8.10 VELOCITY ADDITION

Again consider two inertial systems S and S' moving with relative velocity \mathbf{v} along the xx' axes. Consider a particle at P which is moving with a velocity \mathbf{u} as measured by an observer in S as shown in Fig. 8.4. Let its velocity as measured by an observer in S' be \mathbf{u}' . The velocity components in S and S' are as follows:

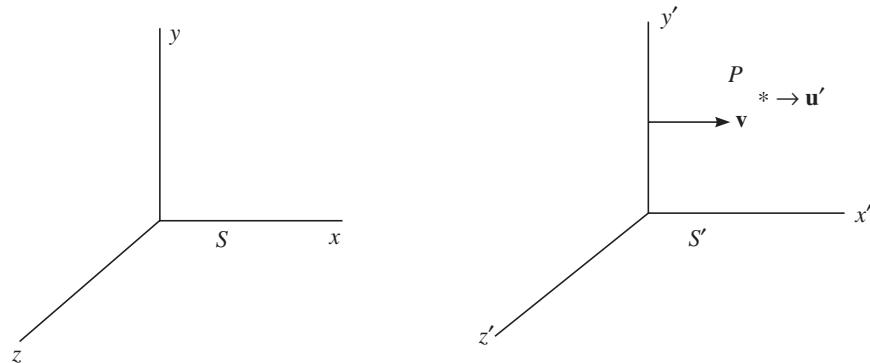


Fig. 8.4 Inertial frames of reference S , S' and event at P .

$$u_x = \frac{dx}{dt}, \quad u_y = \frac{dy}{dt}, \quad u_z = \frac{dz}{dt} \quad (8.20)$$

$$u'_x = \frac{dx'}{dt'}, \quad u'_y = \frac{dy'}{dt'}, \quad u'_z = \frac{dz'}{dt'} \quad (8.21)$$

Differentiation of the Lorentz transformation equation, Eqs. (8.11a–8.11c) gives,

$$\begin{aligned} dx' &= \gamma(dx - \mathbf{v}dt), & dy' &= dy \\ dz' &= dz, \quad dt' = \gamma\left(dt - \frac{\mathbf{v}dx}{c^2}\right) \end{aligned} \quad (8.22)$$

Substituting these values in Eq. (8.21), we get

$$\begin{aligned} u'_x &= \frac{dx - \mathbf{v}dt}{dt - \frac{\mathbf{v}dx}{c^2}} = \frac{dt\left(\frac{dx}{dt} - \mathbf{v}\right)}{dt\left(1 - \frac{\mathbf{v}}{c^2}\frac{dx}{dt}\right)} \\ u'_x &= \frac{u_x - \mathbf{v}}{1 - \left(\frac{\mathbf{v}u_x}{c^2}\right)} \end{aligned} \quad (8.23)$$

$$u'_y = \frac{u_y}{\gamma\left(1 - \frac{\mathbf{v}u_x}{c^2}\right)} = \frac{u_y\sqrt{1 - \beta^2}}{1 - \left(\frac{\mathbf{v}u_x}{c^2}\right)} \quad (8.24)$$

$$u'_z = \frac{u_z}{\gamma\left(1 - \frac{\mathbf{v}u_x}{c^2}\right)} = \frac{u_z\sqrt{1 - \beta^2}}{1 - \left(\frac{\mathbf{v}u_x}{c^2}\right)} \quad (8.25)$$

These are the **Lorentz velocity transformations**. It may be noted that the velocity components u'_y and u'_z also depend on u_x . The inverse transformation is obtained by replacing \mathbf{v} by $-\mathbf{v}$ and interchanging primed and unprimed co-ordinates as follows:

$$u_x = \frac{u'_x + \mathbf{v}}{1 + \left(\frac{\mathbf{v}u'_x}{c^2}\right)} \quad (8.26)$$

$$u_y = \frac{u'_y}{\gamma\left[1 + \left(\frac{\mathbf{v}u'_x}{c^2}\right)\right]} = \frac{u'_y\sqrt{1 - \beta^2}}{1 + \left(\frac{\mathbf{v}u'_x}{c^2}\right)} \quad (8.27)$$

$$u_z = \frac{u'_z}{\gamma \left[1 + \left(\frac{\mathbf{v} u'_x}{c^2} \right) \right]} = \frac{u'_z \sqrt{1 - \beta^2}}{1 + \left(\frac{\mathbf{v} u'_x}{c^2} \right)} \quad (8.28)$$

If u' is along the x' -axis, $u'_x = u'$, $u'_y = 0$, $u'_z = 0$. Then,

$$u' = \frac{u - \mathbf{v}}{1 - \left(u \frac{\mathbf{v}}{c^2} \right)} \quad \text{or} \quad u = \frac{u' + \mathbf{v}}{1 + \left(u' \frac{\mathbf{v}}{c^2} \right)} \quad (8.29)$$

Equation (8.29) is referred to as Einstein's **law of addition of velocities**. Here, \mathbf{v} is the velocity of frame S' with respect to S and u' is the velocity of the event P relative to S' and u is the velocity of the event P relative to S (see Fig. 8.4).

If $u' = c$, the velocity of light will be given as

$$u = \frac{c + \mathbf{v}}{1 + \frac{\mathbf{v}}{c}} = c \quad (8.30)$$

That is, velocity of the source does not add anything to the velocity of light emitted by it. In other words, it is impossible to exceed the velocity of light by adding two or more velocities, no matter how close each of these velocities is to that of light.

8.11 MASS IN RELATIVITY

In Newtonian mechanics, mass is considered to be a constant quantity independent of its velocity. In relativity, like length and time, it is likely to depend on its velocity u . That is, $m = m(u)$ and when $u = 0$, $m = m_0$, the **rest mass** of the particle. We now obtain the form of $m(u)$ by applying the law of conservation of linear momentum, which is a basic principle in physics, together with Lorentz velocity transformations.

Let us consider an inelastic collision between two identical bodies in the inertial system S' which is moving relative to the inertial system S with a velocity \mathbf{v} along the xx' axes (see Fig. 8.5). Let the identical bodies move in opposite directions along the x' axis with velocities u' and $-u'$ in S' . The masses of these bodies as observed from the system S be m_1 and m_2 and their velocities be u_1 and u_2 , respectively. In S' the masses of the bodies are equal and their momenta equal and oppositely directed. Hence, after collision the two bodies will stick together and will be at rest in S' . After collision, the mass $(m_1 + m_2)$ will be moving with velocity \mathbf{v} in system S . Applying the law of conservation of linear momentum to the system S , we obtain

$$m_1 u_1 + m_2 u_2 = (m_1 + m_2) \mathbf{v} \quad (8.31)$$

where, according to Lorentz velocity transformations, u_1 and u_2 are:

$$u_1 = \frac{u' + \mathbf{v}}{1 + \left(u' \frac{\mathbf{v}}{c^2} \right)} = \frac{u' + \mathbf{v}}{1 + x} \quad (8.32)$$

$$u_2 = \frac{-u' + \mathbf{v}}{1 - \left(u' \frac{\mathbf{v}}{c^2} \right)} = \frac{-u' + \mathbf{v}}{1 - x} \quad (8.33)$$

where

$$x = u' \frac{\mathbf{v}}{c^2} \quad (8.34)$$

Substituting the values of u_1 and u_2 in Eq. (8.31), we get

$$m_1 \left(\frac{u' + \mathbf{v}}{1 + x} \right) + m_2 \left(\frac{-u' + \mathbf{v}}{1 - x} \right) = (m_1 + m_2) \mathbf{v}$$

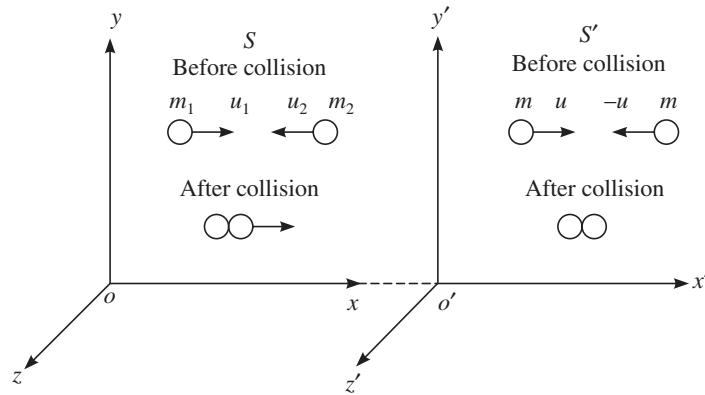


Fig. 8.5 Collision between two bodies in the system S' , as observed from the system S .

$$\begin{aligned} m_1 \left(\frac{u' + \mathbf{v}}{1 + x} \right) - m_1 \mathbf{v} &= m_2 \mathbf{v} - m_2 \left(\frac{-u' + \mathbf{v}}{1 - x} \right) \\ m_1 \left(\frac{u' - \mathbf{v}x}{1 + x} \right) &= m_2 \left(\frac{u' - \mathbf{v}x}{1 - x} \right) \\ \frac{m_1}{m_2} &= \frac{1 + x}{1 - x} \end{aligned} \quad (8.35)$$

Squaring Eq. (8.32) and subtracting c^2 from both sides, we obtain

$$u_1^2 - c^2 = \frac{u'^2 + \mathbf{v}^2 + 2\mathbf{v}u'x}{(1+x)^2} - c^2 = \frac{u'^2 + \mathbf{v}^2 - c^2 - c^2x^2}{(1+x)^2} \quad (8.36)$$

where $u'\mathbf{v}$ is replaced by c^2x . Similarly, from Eq. (8.33), we have

$$u_2^2 - c^2 = \frac{u'^2 + \mathbf{v}^2 - c^2 - c^2x^2}{(1-x)^2} \quad (8.37)$$

From Eqs. (8.36) and (8.37), we obtain

$$\frac{u_2^2 - c^2}{u_1^2 - c^2} = \frac{(1+x)^2}{(1-x)^2} \quad (8.38)$$

$$\frac{1+x}{1-x} = \frac{\sqrt{u_2^2 - c^2}}{\sqrt{u_1^2 - c^2}} = \frac{\sqrt{c^2 \left(1 - \frac{u_2^2}{c^2}\right)}}{\sqrt{c^2 \left(1 - \frac{u_1^2}{c^2}\right)}} = \sqrt{\frac{1 - \frac{u_2^2}{c^2}}{1 - \frac{u_1^2}{c^2}}} \quad (8.39)$$

Combining Eqs. (8.35) and (8.39), we get

$$\frac{m_1}{m_2} = \frac{\sqrt{1 - \frac{u_2^2}{c^2}}}{\sqrt{1 - \frac{u_1^2}{c^2}}} \quad (8.40)$$

If the velocity of the second body as observed with respect to S is zero, then its mass m_2 is m_0 , the rest mass of the identical body. This gives

$$m_1 = \frac{m_0}{\sqrt{1 - \frac{u_1^2}{c^2}}} \quad (8.41)$$

Thus, in general, if a mass is moving with a velocity u relative to an observer, then

$$m = \frac{m_0}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (8.42)$$

In relativity, the invariant quantity is the rest mass m_0 . In classical physics, m_0 is used in place of m also since the speeds acquired by objects are considerably small as compared to that of light. The momentum is defined by

$$\mathbf{p} = m\mathbf{u} = \frac{m_0\mathbf{u}}{\sqrt{1 - \frac{u^2}{c^2}}} \quad (8.43)$$

Newton's second law of motion takes the form:

$$\mathbf{F} = \frac{d(m\mathbf{u})}{dt} = \frac{d}{dt} \left(\frac{m_0\mathbf{u}}{\sqrt{1 - \frac{u^2}{c^2}}} \right) \quad (8.44)$$

Thus, the change in the definition of mass has modified the definition of momentum and Newton's second law of motion.

8.12 MASS AND ENERGY

In this section, let us learn how kinetic energy and total energy get modified in relativity. Let us consider a particle of rest mass m_0 acted upon by a force F through a distance x in time t along the x -axis. Because of the force, the particle attains a final velocity u . Then the kinetic energy T is defined as:

$$\begin{aligned} T &= \int_0^x F dx = \int_0^x \frac{d}{dt} (mu) dx = \int_0^u \frac{dx}{dt} d(mu) \\ &= \int_0^u ud(mu) = m_0 \int_0^u ud(\gamma u) \end{aligned} \quad (8.45)$$

where

$$\gamma = \frac{1}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}}}$$

Carrying out the differentiation, we obtain

$$\begin{aligned} d(\gamma u) &= \gamma du + ud\gamma = \frac{du}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}}} + \frac{u^2}{c^2} \frac{du}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{3}{2}}} \\ &= \left[\frac{c^2 \left(1 - \frac{u^2}{c^2}\right) + u^2}{c^2 \left(1 - \frac{u^2}{c^2}\right)^{\frac{3}{2}}} \right] du \\ &= \frac{du}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{3}{2}}} \end{aligned} \quad (8.46)$$

Substituting this value of $d(\gamma u)$ in Eq. (8.45), we have

$$T = m_0 \int_0^u \frac{udu}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{3}{2}}} = m_0 c^2 \left[\frac{1}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}}} \right]_0^u$$

$$= \frac{m_0 c^2}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}}} - m_0 c^2 = mc^2 - m_0 c^2 \quad (8.47)$$

Eq. (8.47) looks very different from the classical expression $\frac{1}{2} mu^2$. This relation implies that mass is a form of energy. Einstein called $m_0 c^2$, the **rest energy** of the object. It is the total energy of the object measured in a frame of reference in which the object is at rest. By analogy, mc^2 , the sum of kinetic energy and rest energy, is called the **total energy** E . This implies,

$$E = mc^2 \quad (8.48)$$

Eq. (8.48), which states the relationship between mass and energy, is Einstein's **mass energy relation**.

The change of mass to other forms of energy and vice versa have been experimentally confirmed. This interconversion is easily detected in elementary particle physics. Electromagnetic radiation under certain conditions can be converted into electron and positron. The energy produced in nuclear power plants is a result of the loss in mass of the fuel during a fission reaction. Even the radiant energy we receive from the sun is an example of conversion of mass into energy.

8.13 MASSLESS PARTICLES

The mass of a particle varies with velocity according to the relation

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

Multiplying both sides by c^2 and squaring, we get

$$m^2 c^4 \left(1 - \frac{u^2}{c^2}\right) = m_0^2 c^4$$

$$m^2 c^4 = m^2 u^2 c^2 + m_0^2 c^4$$

Using the results $p = mu$ and $E = mc^2$, we get

$$E^2 = c^2 p^2 + m_0^2 c^4 \quad (8.49)$$

This equation allows us to examine the possibility of particles with zero rest mass. With $m_0 = 0$, Eq. (8.49) reduces to

$$E = cp \quad (8.50)$$

That is, when $m_0 = 0$, Eq. (8.50) becomes

$$mc^2 = muc \quad \text{or} \quad u = c$$

Hence, **massless particles may exist, provided they travel at the speed of light.** Such particles cannot be found at rest in any frame, and must have speed c in all inertial frames. Examples of particles in this category are photon, neutrino, and graviton.

SOLVED EXAMPLES

Example 8.1 At what speed does the relativistic value for length differ from the classical value by 1.0 per cent.

Solution: $\frac{L_0 - L}{L_0} = 0.01 \quad \text{or} \quad L = 0.99L_0$

Since, $L = L_0 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} ; \quad \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} = 0.99$

$$\left(1 - \frac{\mathbf{v}^2}{c^2}\right) = 0.98 ; \quad \frac{\mathbf{v}^2}{c^2} = 0.02$$

$$\mathbf{v} = 0.141c$$

Example 8.2 A beam of particles travels at a speed of $0.9 c$. At this speed the mean lifetime as measured in the laboratory frame is 5×10^{-6} s. What is the particle's proper lifetime?

Solution: Proper lifetime $\Delta\tau = \Delta t \sqrt{1 - \frac{\mathbf{v}^2}{c^2}}$

$$\Delta\tau = (5 \times 10^{-6} \text{ s}) \sqrt{1 - (0.9)^2} = 2.18 \times 10^{-6} \text{ s}$$

Example 8.3 Show that $x^2 + y^2 + z^2 - c^2t^2$ is invariant under Lorentz transformation.

Solution: Replacing x, y, z, t by x', y', z', t'

$$\begin{aligned} x^2 + y^2 + z^2 - c^2t^2 &= \gamma^2 (x' + vt')^2 + y'^2 + z'^2 - c^2 \gamma^2 \left[t' + \left(\frac{\mathbf{v}x'}{c^2} \right) \right]^2 \\ &= x'^2 \gamma^2 \left(1 - \frac{\mathbf{v}^2}{c^2} \right) + y'^2 + z'^2 - t'^2 \gamma^2 c^2 \left(1 - \frac{\mathbf{v}^2}{c^2} \right) \\ &= x'^2 + y'^2 + z'^2 - c^2 t'^2 \end{aligned}$$

That is, $x^2 + y^2 + z^2 - c^2t^2$ is invariant under Lorentz transformation.

Example 8.4 A rocket leaves the earth at a speed of $0.6c$. A second rocket leaves the first at a speed of $0.9c$ with respect to the first. Calculate the speed of the second rocket with respect to the earth if (i) it is fired in the same direction as the first one (ii) if it is fired in a direction opposite to the first.

Solution: The first rocket is the reference frame S' moving with a velocity of $0.6c$ relative to S (earth). Hence, $\mathbf{v} = 0.6c$. Second rocket is fired from S' with velocity $0.9c$. Hence, $u' = 0.9c$. Let the velocity of the second rocket relative to S (earth) be u .

$$(i) \quad u = \frac{u' + \mathbf{v}}{1 + \frac{u' \mathbf{v}}{c^2}} = \frac{0.9c + 0.6c}{1 + \frac{(0.9c)(0.6c)}{c^2}} = 0.974c$$

$$(ii) \quad u = \frac{-0.9c + 0.6c}{1 - \frac{(0.9c)(0.6c)}{c^2}} = \frac{-0.3c}{0.46} = -0.652c$$

Example 8.5 The length of a spaceship is measured to be exactly half its proper length. (i) What is the speed of the spaceship relative to the observer on earth? (ii) What is the dilation of the spaceship's unit time?

Solution: (i) Taking the spaceship's frame as S' , the length in the frame S is given by

$$L = L_0 \sqrt{1 - \beta^2} \quad ; \quad \beta = \frac{\mathbf{v}}{c}$$

It is given that $L = \frac{L_0}{2}$. Then,

$$\frac{L_0}{2} = L_0 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} \quad \text{or} \quad 1 - \frac{\mathbf{v}^2}{c^2} = \frac{1}{4}$$

$$\mathbf{v} = 0.866c$$

(ii) Spaceship's unit time means, clock is at rest in S' . From Eq. (8.18), we have

$$\Delta t = \frac{\Delta\tau}{\sqrt{\frac{\mathbf{v}^2}{1 - c^2}}} = \frac{\Delta\tau}{\sqrt{\frac{1}{2}}} = 2\Delta\tau$$

That is, unit time in the S' clock is recorded as twice of unit time by the observer. In other words, the spaceship's clock runs half as fast.

Example 8.6 An inertial frame S' moves with respect to another inertial frame S with a uniform velocity $0.6c$ along the x -axis. The origins of the two systems coincide when $t = t' = 0$. An event occurs at $x_1 = 10$ m, $y_1 = 0$, $z_1 = 0$, $t_1 = 2 \times 10^{-7}$ s. Another event occurs at $x_2 = 40$ m, $y_2 = 0$, $z_2 = 0$, $t_2 = 3 \times 10^{-7}$ s. In S' (i) What is the time difference? (ii) What is the distance between the events?

Solution: (i) From Eq. (8.15), we get

$$\begin{aligned} t'_2 - t'_1 &= \gamma(t_2 - t_1) + \frac{\mathbf{v}}{c^2}(x_1 - x_2) \\ \gamma &= \frac{1}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} = \frac{1}{\sqrt{1 - 0.36}} = \frac{1}{\sqrt{0.64}} = 1.25 \\ t'_2 - t'_1 &= 1.25 \times (3 - 2)10^{-7} \text{ s} + \frac{1.25 \times 0.6(-30 \text{ m})}{3 \times 10^8 \text{ ms}^{-1}} \\ &= 1.25 \times 10^{-7} \text{ s} - 0.75 \times 10^{-7} \text{ s} = 0.5 \times 10^{-7} \text{ s} \end{aligned}$$

(ii) From Lorentz transformation, we get

$$\begin{aligned} x'_1 &= \gamma(x_1 - vt_1); \quad x'_2 = \gamma(x_2 - vt_2) \\ x'_2 - x'_1 &= \gamma(x_2 - x_1) - \gamma v(t_2 - t_1) \\ &= 1.25 \times 30 \text{ m} - 1.25(0.6)(3 \times 10^8 \text{ ms}^{-1}) 10^{-7} \text{ s} \\ &= 37.5 \text{ m} - 22.5 \text{ m} = 15 \text{ m} \end{aligned}$$

Example 8.7 How fast must an unstable particle be moving to travel 20 m before it decays? The mean life time of the particle at rest = 2.6×10^{-8} s.

Solution: The mean life time of 2.6×10^{-8} s is in a frame of reference in which the particle is at rest. That is $\Delta t = 2.6 \times 10^{-8}$ s. Lifetime in the laboratory frame is given by

$$\Delta t = \frac{\Delta\tau}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} = \frac{2.6 \times 10^{-8} \text{ s}}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}}$$

The distance travelled in the laboratory frame in time Δt is 20 m. Hence,

$$\mathbf{v} = \frac{20}{\Delta t} \quad \text{or} \quad \Delta t = \frac{20}{\mathbf{v}} \text{ s}$$

Equating the two values for Δt , we get

$$\frac{20}{\mathbf{v}} \text{ s} = \frac{2.6 \times 10^{-8}}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}} \text{ s}$$

Squaring and simplifying, we obtain

$$\mathbf{v} = 2.8 \times 10^8 \text{ m/s}$$

Example 8.8 The average lifetime of μ -mesons at rest is 2.3×10^{-6} s. A laboratory measurement on μ -meson gives an average lifetime of 6.9×10^{-6} s. (i) What is the speed of the meson in the laboratory? (ii) What is the effective mass of a μ -meson when moving at this speed, if its rest mass is $207m_e$? (iii) What is its kinetic energy?

Solution: (i) Proper time interval $\Delta\tau = 2.3 \times 10^{-6}$ s. For the lifetime in the laboratory Δt we have

$$\Delta t = \frac{\Delta\tau}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{or} \quad 6.9 \times 10^{-6} \text{s} = \frac{2.3 \times 10^{-6} \text{s}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

On solving, we get

$$v = 0.9428 c$$

$$(ii) m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \sqrt{1 - \frac{v^2}{c^2}} = \frac{2.3}{6.9},$$

$$m = 207 m_e \times \frac{6.9}{2.3} = 621 m_e$$

$$(iii) \text{Kinetic energy } T = mc^2 - m_0 c^2 = (621 - 207) m_e c^2 = 414 m_e c^2$$

$$T = 414 (9.1 \times 10^{-31} \text{ kg}) (3 \times 10^8 \text{ ms}^{-1})^2 = 339.07 \times 10^{-13} \text{ J}$$

$$= \frac{339.07 \times 10^{-13} \text{ J}}{1.6 \times 10^{-19} \frac{\text{J}}{\text{eV}}} = 211.9 \times 10^6 \text{ eV} = 211.9 \text{ MeV}$$

Example 8.9 At what speed will the mass of a body be 20% greater than its rest mass?

Solution: It is given that, $\frac{m - m_0}{m_0} = \frac{20}{100}$

$$\frac{m}{m_0} = \frac{120}{100} = \frac{12}{10}$$

We have

$$m = \frac{m_0}{\sqrt{1 - \frac{u^2}{c^2}}} \quad \text{or} \quad \frac{m_0}{m} = \sqrt{1 - \frac{u^2}{c^2}}$$

Combining the two equations

$$\sqrt{1 - \frac{u^2}{c^2}} = \frac{1}{1.2} \quad \text{or} \quad 1 - \frac{u^2}{c^2} = \frac{1}{1.44}$$

$$\frac{u^2}{c^2} = 0.3056 \quad \text{or} \quad u = 0.553 c$$

Example 8.10 If the sun radiates energy at the rate of $4 \times 10^{26} \text{ Js}^{-1}$, evaluate the rate at which its mass is decreasing.

Solution: The change in mass $\Delta m = \frac{\Delta E}{c^2} = \frac{4 \times 10^{26} \text{ Js}^{-1}}{9 \times 10^{16} \text{ m}^2 \text{ s}^{-2}}$

$$\Delta m = 4.44 \times 10^9 \text{ kg s}^{-1}$$

Example 8.11 Find the momentum and velocity of an electron having a kinetic energy 10.0 MeV. The rest energy of electron is 0.512 MeV.

Solution: Total energy $E = T + m_0 c^2 = 10 \text{ MeV} + 0.512 \text{ MeV} = 10.512 \text{ MeV}$

$$E^2 = c^2 p^2 + m_0^2 c^4$$

$$(10.512)^2 \text{ MeV} = c^2 p^2 + (0.512)^2 \text{ MeV}$$

$$cp = 10.5 \text{ MeV} \quad \text{or} \quad p = \frac{10.5}{c} \text{ MeV}$$

Since,

$$E = \frac{m_0 c^2}{\left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}}}, \quad \left(1 - \frac{u^2}{c^2}\right)^{\frac{1}{2}} = \frac{m_0 c^2}{E}$$

$$\frac{u^2}{c^2} = 1 - \left(\frac{m_0 c^2}{E}\right)^2 = 1 - \left(\frac{0.512 \text{ MeV}}{10.512 \text{ MeV}}\right)^2 = 0.9976$$

$$u = 0.9988 c$$

Example 8.12 Reference frame S' moves with a speed \mathbf{v} relative to frame S . Two events occur at the same point in S' at t'_1 and t'_2 . Show that the spatial separation of these events measured in S is $\mathbf{v}(t_2 - t_1)$, where $t_2 - t_1$ is the time interval as measured in S .

Solution: From Eq. (8.12a), we have

$$x_1 = \gamma(x'_1 + \mathbf{v}t'_1), \quad \gamma = \frac{1}{\sqrt{1 - \frac{\mathbf{v}^2}{c^2}}}$$

$$x_2 = \gamma(x'_2 + \mathbf{v}t'_2)$$

$$x_2 - x_1 = \gamma(x'_2 - x'_1) + \gamma\mathbf{v}(t'_2 - t'_1)$$

Since $(x'_2 - x'_1) = 0$

$$x_2 - x_1 = \gamma v(t'_2 - t'_1)$$

As $(t'_2 - t'_1)$ is the proper time interval $\Delta\tau$

$$x_2 - x_1 = \gamma v \Delta\tau = v \Delta t$$

since $\Delta t = \gamma \Delta\tau$ (Eq. 8.18). Expanding Δt , we get

$$x_2 - x_1 = v(t_2 - t_1)$$

Example 8.13 An object has a total energy of 4.5×10^{17} J. The components of its momentum are $p_x = 3.8 \times 10^8$ kg m/s, $p_y = 3 \times 10^8$ kg m/s, and $p_z = 3 \times 10^8$ kg m/s. Calculate its rest mass.

Solution: The momentum of the object p is given by

$$\begin{aligned} p &= \sqrt{p_x^2 + p_y^2 + p_z^2} \\ &= \sqrt{3.8^2 + 3^2 + 3^2} \times 10^8 \text{ kg m/s} \\ &= 5.696 \times 10^8 \text{ kg m/s} \end{aligned}$$

From Eq. (8.49)

$$\begin{aligned} E^2 &= c^2 p^2 + m_0^2 c^4 \\ m_0^2 &= \frac{E^2}{c^4} - \frac{p^2}{c^2} \\ &= \frac{(4.5 \times 10^{17} \text{ J})^2}{(3 \times 10^8 \text{ m/s})^4} - \frac{(5.696 \times 10^8 \text{ kg m/s})^2}{(3 \times 10^8 \text{ m/s})^2} \\ &= 25.00 - 3.6 = 21.4 \text{ kg}^2 \\ m_0 &= 4.63 \text{ kg} \end{aligned}$$

Example 8.14 A high speed probe with a mass $m = 50000$ kg is sent towards a planet at a speed of $u = 0.8c$. What is its momentum as measured by mission control on earth?

Solution: Assuming that the probe travels in a straight line, its momentum p along the direction is given by (Eq. 8.43) as:

$$\begin{aligned} p &= \frac{mu}{\sqrt{1 - \frac{u^2}{c^2}}} \\ &= \frac{(50000 \text{ kg})(0.8 \times 3 \times 10^8 \text{ m/s})}{\sqrt{1 - 0.8^2}} \\ &= 2 \times 10^{13} \text{ kg m/s} \end{aligned}$$

Example 8.15 The speed of an electron in a uniform electric field changes from $u_1 = 0.98 c$ to $u_2 = 0.99 c$. (i) Calculate the change in mass (ii) Compute the work done on the electron to change its velocity. (iii) What is the accelerating potential in volts?

Solution: (i) When velocity is $0.98 c$, the mass becomes

$$m_1 = \frac{m_0}{\sqrt{1 - \frac{u_1^2}{c^2}}} = \frac{m_0}{\sqrt{1 - 0.98^2}} = 5.0 (m_0)$$

When velocity is $0.99 c$, the mass is given by

$$m_2 = \frac{m_0}{\sqrt{1 - 0.99^2}} = 7.1 (m_0)$$

The rest mass of electron $m_0 = 9.11 \times 10^{-31}$ kg

$$\begin{aligned} \text{Change in mass} &= 7.1 m_0 - 5.0 m_0 = (7.1 - 5.0) m_0 \\ &= 2.1 \times 9.11 \times 10^{-31} \text{ kg} \\ &= 19.1 \times 10^{-31} \text{ kg} \end{aligned}$$

(ii) Work done = change in kinetic energy

$$\begin{aligned} &= (\Delta m) c^2 = 2.1 (9.11 \times 10^{-31} \text{ kg}) (3 \times 10^8 \text{ m/s})^2 \\ &= 172.2 \times 10^{-15} \text{ J} \\ &= \frac{172.2 \times 10^{-15} \text{ J}}{1.6 \times 10^{-19} \frac{\text{J}}{\text{eV}}} = 1.08 \text{ MeV} \end{aligned}$$

(iii) Work done = eV , V is the accelerating potential

$$V = \frac{172.2 \times 10^{-15} \text{ J}}{1.6 \times 10^{-19} \text{ C}} = 1.08 \times 10^6 \text{ volts}$$

REVIEW QUESTIONS

- 8.1** What are inertial and noninertial frames of references? Give examples.
- 8.2** Explain the significance of the null result of the Michelson-Morley experiment.
- 8.3** What is a Galilean transformation?
- 8.4** State the postulates of the special theory of relativity.

- 8.5** State and explain the Lorentz transformation equations.
- 8.6** Explain how the length contraction, time dilation, and mass variation expressions might be used to indicate that c is the limiting speed in the universe.
- 8.7** Explain time dilation and length contraction.
- 8.8** State the expressions for the rest energy, kinetic energy, and total energy of a relativistic particle.
- 8.9** What are proper time and proper length?
- 8.10** Show that the addition of a velocity to the velocity of light gives the velocity of light.
- 8.11** “In special theory of relativity mass and energy are equivalent”. Discuss this statement with examples.
- 8.12** Explain how relativity changed our notion about space and time.
- 8.13** Does $E = mc^2$ apply to particles that travel at the speed of light?
- 8.14** “Events that are simultaneous in one reference frame are not simultaneous in another reference frame moving with respect to the first. Comment.
- 8.15** Is mass a conserved quantity in the special theory of relativity?
- 8.16** Show that the velocity of a particle having zero rest mass is equal to the velocity of light.
- 8.17** Draw graphs of energy versus momentum for (i) a particle of zero rest mass
(ii) a particle of non-zero rest mass.

PROBLEMS

1. A rocket travelling away from the earth with a speed of $0.5c$ fires off a second rocket at a speed of $0.6c$ with respect to the first one. Calculate the speed of the second rocket with respect to the earth.
2. Two spaceships leave the earth in opposite directions, each with a speed of $0.5c$ with respect to the earth. What is the velocity of spaceship 1 relative to spaceship 2?
3. At what speed does the relativistic value for time differ from the classical value by 2 per cent?
4. An object passes at a speed of $0.8c$. Its length is measured to be 72.5 m. What would be its length at rest?
5. A free neutron has an average lifetime of 1000 s. How fast must a beam of neutrons be travelling for them to have a lifetime twice this long with respect to the laboratory value?
6. A proton has a kinetic energy of m_0c^2 . Find its momentum in units of MeV/c.
7. A particle with a mean life time of 10^{-6} s moves through the laboratory at a speed of $0.8c$. What will be its lifetime as measured by an observer in the laboratory?
8. If the kinetic energy of an electron is 5 MeV, what is its velocity?

9. Calculate the rest energy in MeV of an electron and a proton. Mass of electron- 9.11×10^{-31} kg, Mass of proton = 1.67×10^{-27} kg.
10. What is the speed of a beam of particles if their mean lifetime is 3×10^{-7} s? Their proper lifetime is 2.6×10^{-7} s.
11. A meson having a mass of 2.4×10^{-28} kg travels at a speed of $v = 0.8c$. What is its kinetic energy?
12. At what speed will the mass of a body be twice its rest mass?
13. π -mesons coming out of an accelerator have a velocity of $0.99c$. If they have a mean lifetime of 2.6×10^{-8} s in the rest frame, how far do they travel before decay?
14. Calculate the mass of a particle whose kinetic energy is half its total energy. How fast is it travelling?
15. Evaluate the speed of a particle when its kinetic energy equals its rest energy.

C H A P T E R

9

Quantum Mechanics

Classical physics treated particles and waves as distinct entities. A wave distributes its energy throughout the space in wavefronts, whereas in a particle, the energy is confined to a small region of space. However, according to quantum mechanics, radiation and material particles have dual nature, wave and particle. To start with, we shall discuss the wave particle duality exhibited by radiation and matter, followed by the uncertainty principle which is a consequence of the wave nature of matter. Based on the wave nature of matter, we develop some of the basics of quantum mechanics including the Schrodinger equation for a single particle and the probability interpretation of the wave function. A more systematic presentation of the formulation will be done based on a set of postulates. Solution of the Schrodinger equation for a particle in a one-dimensional box and the concept of quantum mechanical tunneling will also be discussed.

9.1 BREAKDOWN OF CLASSICAL PHYSICS

By the end of the 19th century, classical physics appeared capable of explaining all observable phenomena in physics. Then came some of the experimental discoveries of tremendous significance. In 1895, Roentgen discovered X-rays, in 1896, Becquerel discovered radioactivity, and in 1897, Thomson discovered electron. These phenomena are completely different from anything observed before. Along with these discoveries, a large amount of experimental observations starting with blackbody radiation and optical spectra accumulated which the classical theory failed to explain.

9.1.1 Quantization of Matter and Charge

The idea that matter is composed of tiny particles or atoms dates back to the Greek philosopher Democritus. A theory of continuous matter could hardly account for many of the observed phenomena. It was the work of Gassendi, Dalton, Avogadro, Robert Brown and others that led to the general acceptance of the atomic theory of matter. Thus, matter is not continuous as it appears, but is discrete or quantized.

Another great quantization discovery is that of electric charge. The first estimate of the order of electric charge was obtained by Faraday, in 1833. Later, it was confirmed by the experiments of Zeeman (1896), J.J.Thomson (1897), and Townsend (1897). All experiments pointed to the fact that electric charge (e) is quantized. The accepted value of e today is $1.60218 \times 10^{-19} C$.

9.1.2 Photoelectric Effect

In 1887, Hertz discovered that ultraviolet rays falling on certain metals eject electrons from the metal surface. An explanation for this phenomenon was attempted on the basis of classical ideas, which suggested that light of high intensity would eject electrons with greater kinetic energy. However, experiments showed that the kinetic energy of the ejected electrons is independent of the intensity of the incident radiations. It was found to depend on the frequency of radiation. There was even a certain critical frequency below which there was no emission of photoelectrons. On classical ground, these ideas stood unexplained.

9.1.3 Specific Heat

The molar specific heat at constant volume for monoatomic and diatomic gases based on the law of equipartition of energy, were found to be $3R/2$ and $5R/2$ respectively. The agreement was excellent with the experimental results for monoatomic gases, whereas for diatomic gases, the observed specific heats differed for different gases and increased with temperature. For polyatomic gases, the discrepancy was more pronounced. In the case of solids too, the specific heat was found to depend on temperature.

9.1.4 Optical Spectra

In the later half of the 19th century, it was known that the spectrum of hydrogen consisted of lines with certain regularities. It was J.J. Balmer who in 1985 first discovered that the spectral lines of hydrogen in the visible region could be fitted into an empirical formula. Later, other spectral series were found to exist in the infrared and ultraviolet regions. Classical physics failed to give the correct explanation for these phenomena.

9.1.5 Blackbody Radiation Curves

The spectral energy density u_v of blackbody radiation from a blackbody cavity depends only on frequency v and temperature T of the cavity. Based on thermodynamic arguments, Wien attempted to explain the observed spectral energy density u_v versus frequency v curves. The agreement with experimental results was good only in the high frequency region (see Fig. 9.1). Treating Radiation inside a blackbody as standing electromagnetic waves, Rayleigh and Jeans estimated the number of modes of vibration per unit volume in the frequency range v and $v + dv$ as

$$\frac{8\pi v^2 dv}{c^3}$$

They evaluated the spectral energy density u_v by taking its product with the average energy kT of an oscillator of frequency v . Their expression agreed with experiment only at low frequencies. The methods based on classical theories thus failed to give a single formula that could agree with the experimental blackbody radiation curve for the entire frequency range. No wonder this disagreement led to a complete revision of our ideas of physics.

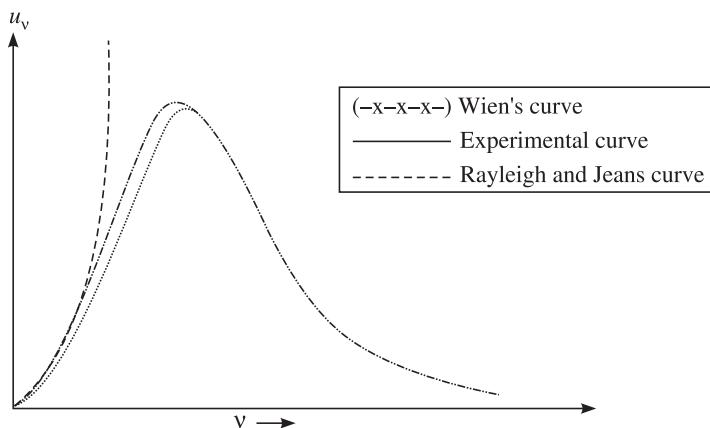


Fig. 9.1 Blackbody radiation curve.

Thus, by 1900, physicists were convinced that all was not well with classical physics as had been assumed. Already things were there to indicate that revolutionary ideas were in the offing. The year 1900 can conveniently be taken as the turning point which marked a real division between the classical era and the new era in physics.

9.2 PLANCK'S QUANTUM THEORY

9.2.1 Planck's Hypothesis

The problem that confronted Max Planck was a theoretical explanation for the blackbody radiation curves. In 1900, Planck modified Wien's formula in such a way that it fitted with the experimental curves precisely, and then he looked for a sound theoretical basis for the formula. He assumed that the atoms of the walls of the blackbody behave like tiny electromagnetic oscillators, each with a characteristic frequency of oscillation. The oscillators emit electromagnetic energy into the cavity and absorb electromagnetic energy from it. Planck then boldly put forth the following suggestions regarding the atomic oscillators:

- (i) An oscillator can only have energies given by

$$E_n = nhv, \quad n = 0, 1, 2, 3, \dots \quad (9.1)$$

where v is the oscillator frequency and h is a constant known as **Planck's constant**. Its value is 6.626×10^{-34} Js. In other words, the oscillator energy is quantized.

(ii) Oscillators can absorb or emit energy only in discrete units called **quanta**.
That is,

$$\Delta E_n = \Delta nh\nu = h\nu \quad (9.1a)$$

An oscillator, in a quantized state, neither emits nor absorbs energy.

9.2.2 Average Energy of an Oscillator

The average energy of an oscillator ε can be evaluated on the basis of classical distribution formula as follows:

$$\begin{aligned} \varepsilon &= \frac{\sum_{n=0}^{\infty} E_n e^{-E_n/kT}}{\sum_{n=0}^{\infty} e^{-E_n/kT}} = \frac{\sum_{n=0}^{\infty} nh\nu e^{-nh\nu/kT}}{\sum_{n=0}^{\infty} e^{-nh\nu/kT}} \\ &= \frac{h\nu e^{-h\nu/kT} + 2h\nu e^{-2h\nu/kT} + \dots}{1 + e^{-h\nu/kT} + e^{-2h\nu/kT} + \dots} \end{aligned}$$

Writing $e^{-h\nu/kT} = x$, we get

$$\begin{aligned} \varepsilon &= \frac{h\nu x(1 + 2x + 3x^2 + \dots)}{1 + x + x^2 + \dots} = \frac{h\nu x(1 - x)^{-2}}{(1 - x)^{-1}} \\ &= \frac{h\nu x(1 - x)^{-2}}{(1 - x)(1 - x)^{-2}} = \frac{h\nu x}{1 - x} = \frac{h\nu e^{-h\nu/kT}}{1 - e^{-h\nu/kT}} \\ &= \frac{h\nu}{e^{h\nu/kT} - 1} \quad (9.2) \end{aligned}$$

9.2.3 Planck's Radiation Formula

The quantization condition, Eq. (9.1), thus invalidates the theorem of equipartition of energy, which is based on classical physics. It is known that the number of oscillators per unit volume in the frequency range ν and $\nu + d\nu$ is $(8\pi\nu^2/c^3)d\nu$. With the above expression for average energy ε , the spectral energy density u_ν is given by

$$u_\nu = \frac{8\pi h\nu^3}{c^3} \frac{d\nu}{e^{h\nu/kT} - 1} \quad (9.2a)$$

Equation (9.2a) is the Planck's radiation formula, which reduces to the Wien's or Rayleigh-Jeans' law according as $\left(\frac{hv}{kT}\right) \gg 1$ or $\left(\frac{hv}{kT}\right) \ll 1$.

Planck's explanation of the blackbody radiation curves in 1900 provided the crucial step in the development of quantum ideas in physics which was put on a firm basis by the pioneering work of Einstein, Bohr, and others in later years. The concept that energy is quantized was so radical that Planck himself was reluctant to accept it. This can be seen from his own words "my futile attempts to fit the elementary quantum of action (that is the quantity h) somehow into the classical theory continued for a number of years and they cost me a great deal of effort".

9.3 DUAL NATURE OF RADIATION

In an attempt to understand the nature of light radiation, Sir Isaac Newton (1680) formulated the corpuscular theory of light. It failed to explain phenomena like interference, diffraction, and polarization. In the mean time Huygens proposed the wave theory of light, which got established as it explained most of the experimental results. In 1864, C. Maxwell proposed his electromagnetic theory in the form of four equations, called **Maxwell's equations**. He showed that all radiations are electromagnetic in nature and have the same speed in free space. He also proved that these radiations proceed in the form of waves with their electric and magnetic fields in planes perpendicular to each other and to the direction of propagation. The theory got established when Hertz (1887) produced electromagnetic waves in the laboratory. The particle nature of radiation was again revived with the advent of Planck's quantum hypothesis which was put on a firm basis by the pioneering work of Einstein, Compton and others.

Thus, certain phenomena can only be explained by assigning a particle nature to radiation. However, the wave nature of radiation is needed to explain phenomena such as interference, diffraction, and polarization. We can therefore arrive at a conclusion that light radiation possesses a dual nature-particle and wave. At times it seems to behave like a wave and at certain other times it seems to behave like a particle. In other words, both the aspects are needed for a complete description of physical phenomena. Hence, the wave and particle aspects are complementary.

9.4 EARLY MODELS OF THE ATOM

With the discovery of electron, scientists began to suggest different models for the atom. Of these, the most important ones are the Rutherford's nuclear model and the Bohr model of the hydrogen atom. In the Rutherford model, the atom consists of a tiny, but massive positively charged nucleus surrounded by electrons some distance away, moving in orbits around the nucleus. Though this model was a major step in the right direction, it was not a stable one, as the electrons were expected to spiral into the nucleus. It was Niels Bohr (1913) who provided the necessary modification to the Rutherford model by incorporating the newly developed quantum ideas. His theory was based on the following postulates:

- (i) The electrons move about the nucleus in certain circular orbits without radiating energy. He called these orbits as **stationary states**.

- (ii) The atom radiates energy when the electron makes a transition from one stationary state to another and the frequency ν of the emitted radiation is related to the energy of the orbits by the relation

$$h\nu = E_i - E_f \quad (9.3)$$

where h is the Planck's constant, E_i and E_f are the energies of the initial and final states.

- (iii) Angular momentum is quantized; it can assume only the values $n\hbar/2\pi$, where n is an integer:

$$L = mvr = n \frac{\hbar}{2\pi} = n\hbar, \quad n = 1, 2, 3, \dots \quad (9.4)$$

where the integer n is called the **quantum number** and $\hbar = h/2\pi$ is called the **modified Planck's constant**. Equation (9.4) is referred to as **Bohr's quantum condition**. Equation (9.3) is referred to as **Bohr frequency condition**. Based on these postulates, Bohr derived expressions for the radius of the orbits and the total energy E_n of the electron in an orbit.

$$E_n = -\frac{mk^2e^4}{2\hbar^2n^2}, \quad k = \frac{1}{4\pi\epsilon_0}, \quad n = 1, 2, 3, \dots \quad (9.5)$$

where ϵ_0 is the permittivity of vacuum and has the value $8.984 \times 10^9 \text{ Nm}^2 \text{ C}^{-2}$. Bohr's theory was successful in explaining many of the observed experimental results. It was largely adhoc in nature. Bohr's reason for Eq. (9.4) was simply that it worked. Ten years later a reason was found by Louis de Broglie of France.

9.5 MATTER WAVE

We noted that the classic experiments on interference, diffraction, and polarization showed that the wave theory of light rests on firm experimental basis. The quantum theory developed by Planck and Einstein, and the experimental observations on photoelectric effect and Compton effect have placed the particle theory of light also on a firm basis. Hence, physicists accepted the wave-particle duality of light radiation as a reality.

9.5.1 De Broglie Hypothesis

In 1923, Louis de Broglie extended the concept of wave-particle duality exhibited by radiation to matter, in particular, to electrons. He was influenced by the grand symmetry of nature since the universe consists entirely of matter and radiation. He reasoned that if light can be particle like at times, then matter can also be wave at times. Therefore, he concluded that if light has a wave-particle duality, the other component of the universe, matter, might also possess the wave-particle duality. Another reason in favour of the wave nature of matter was that it provided an explanation for the angular momentum quantization of Bohr, as given in Eq. (9.4). The wave associated with a material particle in motion is called a **matter wave**.

9.5.2 Wave Packet

The wave associated with a particle in motion can only be represented by a wave disturbance that has a large amplitude in the vicinity of the particle and a small amplitude in regions where we do not expect to find the particle. In other words, it must be represented by a wave that is confined to a small region of space around the particle. Such a wave is called a **wave packet**. An ideal wave packet is represented in Fig. 9.2. The wave packet moves with the velocity of the particle.

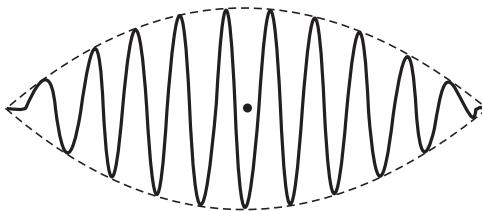


Fig. 9.2 An ideal wave packet.

9.6 DE BROGLIE WAVELENGTH

Photons are particles of zero rest mass. From relativity, for zero rest mass particles, we have

$$E = cp \quad (9.6)$$

where c is the velocity of the particle which is the velocity of light and p is its momentum. For radiation, we have

$$E = h\nu = \hbar\omega \quad (9.7)$$

Combining the two relations, we get

$$p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda} \quad \text{or} \quad \lambda = \frac{h}{p} \quad (9.8)$$

De Broglie proposed that the wavelength and frequency of the matter wave would be related respectively to the momentum and energy by the same relationships as for a photon. Thus, for a particle of mass m and total energy E travelling with speed v , the wavelength λ and frequency ν are given by

$$\lambda = \frac{h}{p} = \frac{h}{mv} \quad (9.9)$$

and $\nu = \frac{E}{h}$ (9.10)

These are referred to as **de Broglie relations**. The wavelength λ is called the **de Broglie wavelength** of the particle.

For an electron accelerated from rest through a potential difference of V volts, assuming non-relativistic mechanics, we have

$$\frac{1}{2}mv^2 = eV \quad \text{or} \quad v = \sqrt{\frac{2eV}{m}}$$

With this value of v , Eq. (9.9) reduces to

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2meV}} \quad (9.11)$$

Substituting the values of constants, for electrons

$$\lambda = \sqrt{\frac{150}{V}} \times 10^{-10} \text{ m} = \sqrt{\frac{150}{V}} \text{ Å} \quad (9.12)$$

One of the de Broglie's original arguments in favour of the wave nature of electron was that it could give Bohr's quantization of angular momentum, Eq. (9.4). Davisson and Germer (1927) and G.P. Thomson (1927) measured the wavelength of matter waves associated with electrons in motion and confirmed de Broglie's hypothesis.

9.7 THE DAVISSON-GERMER EXPERIMENT

The de Broglie wavelength associated with low energy electrons is of the order of the atomic spacing of most of the crystalline solids. Therefore, atoms of a crystal can serve as a three-dimensional array of diffracting centres for the matter waves. Figure 9.3(a) shows a schematic representation of the apparatus used by Davisson and Germer.

Electrons from a heated filament is accelerated through a potential difference V in the electron gun and allowed to strike a single crystal of nickel. The detector can be moved to any angle ϕ relative to the incident beam. Davisson and Germer measured the intensity of the scattered electrons as a function of the angle ϕ . Figure 9.3(b) shows the results when

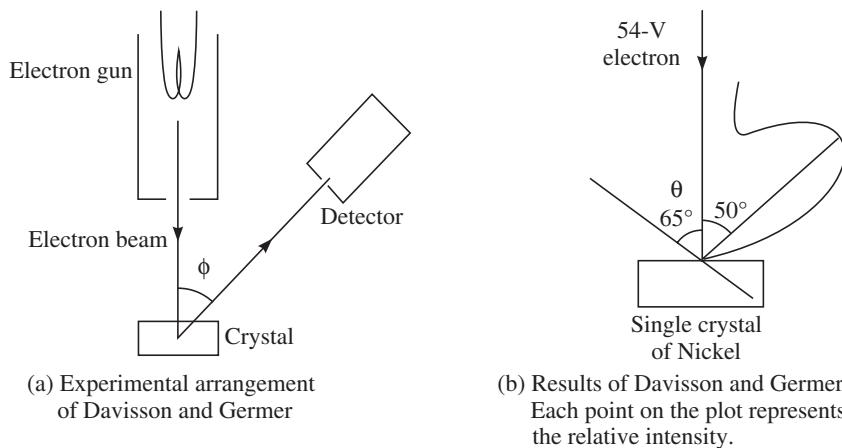


Fig. 9.3 The Davisson-Germer Experiment.

the accelerating voltage is 54V. There is an intense reflection of the beam at the angle $\phi = 50^\circ$. Such reflections can be accounted for by assuming that the electron beam has a wave associated

with it and that Bragg reflections occur from certain families of atomic planes (Section 6.13). The situation is similar to using a reflection grating for light. Fig. 4.2 shows such a reflection satisfying the Bragg condition

$$2d \sin \theta = n\lambda \quad ; \quad n = 1, 2, 3, \dots \quad (9.13)$$

where d refers to the separation between crystal planes and $n = 1, 2, 3, \dots$ refer to first order, second order, etc., respectively. For Nickel, $d = 0.91 \text{ \AA}$ and $\theta = 65^\circ$. From Eq. (9.13) λ is found to be 1.65 \AA , which is in agreement with the calculated value of 1.66 \AA .

9.8 THOMSON'S EXPERIMENT

The concept of wave nature of material particles was independently tested by G.P.Thomson also G.P.Thomson's experimental arrangement is analogous to Laue's X-ray diffraction method. The arrangement shown in Fig. 9.4(a) consists of a discharge tube in which a beam of electrons from a cathode C is accelerated by a potential difference ranging from 10,000 to 50,000 volts. The electrons collimated by the tube A fall on a thin gold film of thickness of the order of 10^{-6} cm . The apparatus is evacuated to avoid collision of electrons with the molecules of the gas. The diffracted beam is allowed to fall on a fluorescent screen S or on a photographic plate P . The photograph of the diffracted beam shows a system of concentric rings [see Fig. 9.4(b)]. Measuring the radii of the rings and the distance between the film and the photographic plate, the angle of diffraction θ can be obtained. Knowing the distance between the atomic planes d , the wavelength of the diffracted beam can be calculated. The experiment clearly demonstrates the wave nature of electron as diffraction pattern can only be produced by waves.

In the acceleration process the electron behaves like a particle, in the diffraction process it behaves like a wave, and in the detector it behaves like a particle. The electron that showed wave aspect in one part of the experiment showed particle aspect in two other parts of the same experiment. That is, for a complete description of physical phenomena, both particle and wave aspects of material particles are required. Hence, the new theory which we are looking for must account for the dual nature of radiation and matter.

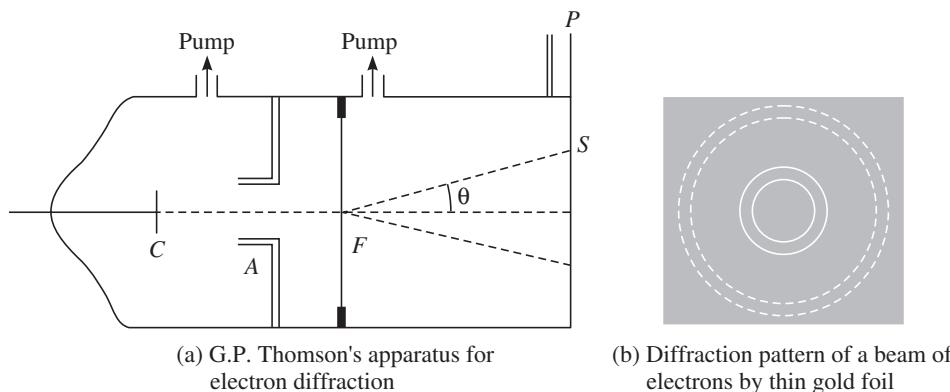


Fig. 9.4 Thomson's experiment.

9.9 UNCERTAINTY PRINCIPLE

In any measurement, there is always some uncertainty or error involved, though it can be decreased by using more precise instruments. According to quantum mechanics, there is a limit to the accuracy of certain measurements. This limit is not due to the measuring device, but is something inherent in nature. It is a consequence of two factors, the wave nature of matter and the interaction between the object observed and the observing device. Werner Heisenberg critically analysed this uncertainty in measurements and formulated the uncertainty principle in 1927. It states that we cannot measure both the position and the momentum of an object precisely at the same time. If Δx and Δp_x are the uncertainties in the co-ordinate x and the momentum p_x , the uncertainty principle states that

$$(\Delta x)(\Delta p_x) \geq h \quad (9.14)$$

In three dimensions, in addition to Eq. (9.14), for y and z components, we have

$$(\Delta y)(\Delta p_y) \geq h \quad (9.14a)$$

$$(\Delta z)(\Delta p_z) \geq h \quad (9.14b)$$

As the product of uncertainties is a universal constant, the more precisely we determine one variable, the less accurate is our determination of the other variable. The uncertainty (ΔA) in a dynamical variable A is usually defined as the root mean square deviation from the mean. That is,

$$(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle (A^2 - 2A\langle A \rangle + \langle A \rangle^2) \rangle = \langle A^2 \rangle - \langle A \rangle^2 \quad (9.15)$$

With this definition, a more accurate calculation gives

$$(\Delta x)(\Delta p) \geq \frac{\hbar}{2} \quad (9.16)$$

This is the more accurate uncertainty relation.

Another useful form of the uncertainty principle relates energy and time:

$$(\Delta E)(\Delta t) \geq \frac{\hbar}{2} \quad (9.17)$$

This form of uncertainty tells us that the energy of an object can be uncertain or may even be non-conserved by an amount ΔE for a time $\Delta t \equiv \hbar/(2\Delta E)$

The consequences of the uncertainty principle is that we can only talk of the probability of getting a particular value for a variable. This limit of accuracy in measurement is because of the wave nature of electron. The product of uncertainties in co-ordinate and momentum is independent of the particle mass, wavelength of light used, etc. Hence, it is a fundamental law of nature. As a deterministic model in the classical sense has no basis, a new mechanics is needed which must allow only for probable values for dynamical variables. The new mechanics was formulated independently by Heisenberg (1925) and Schrödinger (1926). Both are identical and the name quantum mechanics is now generally used for the new mechanics.

9.10 ABSENCE OF ELECTRON IN THE NUCLEUS

From scattering experiments, it is known that the nuclear diameter is of the order of 10^{-15}m . If an electron exists inside the nucleus, the uncertainty in its position must be of this order. That is, $\Delta x \cong 10^{-15}\text{m}$. Substituting this value of Δx in Eq. (9.16), we get

$$\Delta p \cong \frac{\hbar}{2 \times 10^{-15} \text{ m}} \cong \frac{6.626 \times 10^{-34} \text{ Js}}{4(3.143)(10^{-15} \text{ m})} \cong 5.27 \times 10^{-20} \text{ kg m/s}$$

Then the momentum of the electron p must at least be of the order of:

$$p \cong 5.27 \times 10^{-20} \text{ kg m/s} \quad (9.18)$$

When the energy of the electron is very large as compared to its rest energy, then

$$\begin{aligned} E \cong cp &= (3 \times 10^8 \text{ m/s})(5.27 \times 10^{-20} \text{ kg m/s}) = 15.81 \times 10^{-12} \text{ J} \\ &= \frac{15.81 \times 10^{-12} \text{ J}}{1.6 \times 10^{-19} \frac{\text{J}}{\text{eV}}} = 98.8 \times 10^6 \text{ eV} = 98.8 \text{ MeV} \end{aligned} \quad (9.19)$$

For an electron to be a constituent of the nucleus it should have an energy of this order. However, the energy of an electron emitted in a β -decay experiment is of the order of only a few MeV. Hence, one does not expect the electron to be a constituent of the nucleus.

9.11 UNCERTAINTY IN FREQUENCY OF LIGHT EMITTED BY AN ATOM

Equation (9.17) implies that the energy of a system can be measured exactly, only if an infinite amount of time is available for the measurement. If an atom is in an excited state, it does not remain there indefinitely, but makes transition to a lower energy state. We can take the mean time for decay τ , called the life time, as a measure of the time available to determine the energy of the state. Hence, the uncertainty in time is of the order of τ . From Eq. (9.17), we have

$$\Delta E \cong \frac{\hbar}{2\Delta t} = \frac{\hbar}{2\tau} \quad (9.20)$$

Since $E = hv$, $\Delta E = h\Delta v$. Substituting this value of ΔE in Eq. (9.20), we get

$$\Delta v \cong \frac{1}{4\pi\tau} \quad (9.21)$$

Hence, the frequency of light emitted by an atom is uncertain by this amount. For atomic systems, $\tau \sim 10^{-8}\text{s}$. Hence, $\Delta v \sim 10^7 \text{ Hz}$. This width of the emitted line is experimentally observed when the pressure of the emitting gas is low.

9.12 THE WAVE FUNCTION

Waves on a string are described by the displacement $y(x, t)$ of the string. In the case of a sound wave in air, the pressure $p(x, t)$ varies in space and time. In electromagnetic waves, the fields,

E and B , are the ones that vary in space and time. Therefore, to characterize the de Broglie wave associated with a material particle, we require a quantity that varies in space and time. The variable quantity is called the **wave function** for the particle and is usually designated by Ψ which is a function of the co-ordinates (x, y, z) and time t .

The displacement of a wave can either be positive or negative. In analogy, the wave function $\Psi(x, y, z, t)$ can have positive as well as negative values. The uncertainty principle tells us that we can only get the probability of finding the particle at (x, y, z) at time t . As probability cannot be negative, $\Psi(x, y, z, t)$ cannot be a direct measure of the presence of the particle. Hence, Ψ as such is not observable. But it must in some way indicate the presence of the particle, as it represents the wave associated with the particle in motion. For a particle having a well-defined momentum, the uncertainty in co-ordinate is infinite. Hence, the wave associated with it will be of infinite extent. Then, a free particle moving along x -axis with a definite momentum is described by the plane wave

$$\Psi(x, t) = A e^{i(kx - \omega t)} \quad (9.22)$$

where $k = \frac{2\pi}{\lambda}$ is the **propagation constant** and ω is the angular frequency. In three dimensions

$$\Psi(\mathbf{r}, t) = A e^{i(kr - \omega t)} \quad (9.23)$$

Such particles are non-localized particles. For a localized particle, the wave function would be represented by the wave packet.

9.13 TIME DEPENDENT SCHRÖDINGER EQUATION

Classical theories of particles and waves are based on equations of motion. In quantum mechanics the state of a particle is described by the wave function $\Psi(\mathbf{r}, t)$. To study the motion of systems in a general way, one requires an equation of motion, which was derived by Erwin Schrödinger in 1926. This equation, called the **Schrödinger equation**, governs the propagation of matter waves.

Experiments have proved beyond doubt that particles exhibit wavelike properties and the wavelength of the matter wave is given by Eq. (9.9) as:

$$\lambda = \frac{\hbar}{p} \quad (9.24)$$

The energy E of the particle is related to the frequency through the relation Eq. (9.10) as follows:

$$E = h\nu = \hbar\omega \quad (9.25)$$

9.13.1 Free Particle Wave Equation in One Dimension

A free particle moving along x -axis with a definite momentum (a non-localized particle) is described by the infinite plane wave (Eq. 9.22):

$$\Psi(x, t) = A \exp[i(kx - \omega t)] \quad (9.26)$$

Using Eqs. (9.24) and (9.25), we get

$$k = \frac{2\pi}{\lambda} = \frac{2\pi p}{\hbar} = \frac{p}{\hbar} \quad \text{and} \quad \omega = \frac{E}{\hbar} \quad (9.27)$$

Substituting these values in Eq. (9.26), we get

$$\Psi(x, t) = A \exp\left[\frac{i}{\hbar}(px - Et)\right] \quad (9.28)$$

Differentiating Eq. (9.28) with respect to x , we get

$$\frac{\partial \Psi}{\partial x} = \left(\frac{ip}{\hbar}\right) A \exp\left[\frac{i}{\hbar}(px - Et)\right] = \frac{ip}{\hbar} \Psi$$

Multiplying both sides by $-i\hbar$, we obtain

$$-i\hbar \frac{\partial \Psi}{\partial x} = p\Psi \quad (9.29)$$

On differentiation, we get

$$\frac{\partial^2 \Psi}{\partial x^2} = \left(\frac{ip}{\hbar}\right)^2 A \exp\left[\frac{i}{\hbar}(px - Et)\right] = -\frac{p^2}{\hbar^2} \Psi$$

Multiplying both sides by $-\frac{\hbar^2}{2m}$, we obtain

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} = \frac{p^2}{2m} \Psi \quad (9.30)$$

Differentiation of Eq. (9.28) with respect to t gives

$$\frac{\partial \Psi}{\partial t} = \left(-\frac{iE}{\hbar}\right) A \exp\left[\frac{i}{\hbar}(px - Et)\right] = -\frac{iE}{\hbar} \Psi$$

Multiplying both sides by $i\hbar$, we get

$$i\hbar \frac{\partial \Psi}{\partial t} = E\Psi \quad (9.31)$$

The right hand side of Eqs. (9.30) and (9.31) are equal, since for a classical particle

$$E = \frac{p^2}{2m} \quad (9.32)$$

Hence, equating the left hand sides, we get

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} \quad (9.33)$$

Equation (9.33) is the one dimensional time dependent Schrödinger equation for a free particle. In three dimensions, the equation becomes

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} \right) = -\frac{\hbar^2}{2m} \nabla^2 \Psi \quad (9.34)$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (9.35)$$

9.13.2 Energy and Momentum Operators

Equations (9.31) and (9.29) can be written in the form

$$\left(i\hbar \frac{\partial}{\partial t} \right) \Psi = E\Psi \quad \text{and} \quad \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi = p\Psi$$

This suggests that the energy E and the momentum p can be considered as differential operators.

$$E \rightarrow i\hbar \frac{\partial}{\partial t} \quad \text{and} \quad p \rightarrow -i\hbar \frac{\partial}{\partial x} \quad (9.36)$$

In three dimensions, we have

$$E \rightarrow i\hbar \frac{\partial}{\partial t} \quad \text{and} \quad \mathbf{p} = -i\hbar \nabla \quad (9.37)$$

9.13.3 General Wave Equation

The energy E (also called Hamiltonian) of a particle in a potential $V(r, t)$ is given by

$$E = \frac{p^2}{2m} + V(r, t) \quad (9.38)$$

Schrödinger guessed that the operator for position co-ordinate is r and that for time is t itself. Hence, the operator for the potential is $V(r, t)$. Substituting the operators for E , \mathbf{p} , and $V(r, t)$ in Eq. (9.38) and allowing the resulting operator equation to operate on the wave function $\Psi(\mathbf{r}, t)$, we get

$$i\hbar \frac{\partial \Psi}{\partial t} = \left[-\frac{\hbar^2}{2m} \nabla^2 + V(r, t) \right] \Psi(\mathbf{r}, t) \quad (9.39)$$

Equation (9.39) is the **time-dependent Schrödinger equation** for a particle moving in a potential $V(r, t)$. The presence of i in the equation makes the solution of Ψ , in general, a complex quantity.

9.14 TIME INDEPENDENT SCHRÖDINGER EQUATION

In many cases of interest, the potential V is a function of the co-ordinates only. In such cases, we can separate the variables in Eq. (9.39) and obtain two equations, one depending on the variable t and the other on the variable \mathbf{r} . We can write

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r}) \phi(t) \quad (9.40)$$

where $\psi(\mathbf{r})$ is a function of space co-ordinates and $\phi(t)$ is a function of time only. Substituting Eq. (9.40) in Eq. (9.39), we get

$$\psi(\mathbf{r}) \cdot i\hbar \frac{d\phi(t)}{dt} = \phi(t) \left[-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right] \psi(\mathbf{r})$$

Dividing both sides by $\psi\phi$, we get

$$i\hbar \frac{1}{\phi} \frac{d\phi}{dt} = \frac{1}{\psi} \left[-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right] \psi(\mathbf{r}) \quad (9.41)$$

The left side of this equation is a function of t only, while the right side is a function of \mathbf{r} only. This is possible only if both sides are equal to a constant, say E . Equating each side to E , we get

$$i\hbar \frac{d\phi(t)}{dt} = E\phi(t) \quad (9.42)$$

and
$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (9.43)$$

Equation (9.43) is the **time independent Schrödinger equation** or simply the **Schrödinger equation**. It will be shown that the constant E is the total energy of the system (Section 9.19).

9.15 PROBABILITY INTERPRETATION OF THE WAVE FUNCTION

We shall now discuss the probability interpretation of the wave function. In Section 9.12, we noticed that the wave function $\Psi(\mathbf{r}, t)$ is not observable. But it is certain that Ψ must in some way be an index of the presence of the particle at (\mathbf{r}, t) . A statistical interpretation of the wave function, now universally accepted, was suggested by Born in 1926. According to him, the quantity $\Psi^* \Psi = |\Psi|^2$ is the **probability density** $P(\mathbf{r}, t)$:

$$P(\mathbf{r}, t) = |\Psi(\mathbf{r}, t)|^2 \quad (9.44)$$

The probability of finding the system in a volume element $d\tau$ is then $|\Psi|^2 d\tau$. This quantity when integrated from $-\infty$ to $+\infty$ must then be the total probability, which is equal to unity. That is,

$$\int_{-\infty}^{\infty} |\Psi|^2 d\tau = 1 \quad (9.45)$$

Another important parameter concerning the probability interpretation of the wave function is the **probability current density** $\mathbf{j}(\mathbf{r}, t)$, which is the flow of probability across unit area in unit time. It is a vector defined by the relation:

$$\mathbf{j}(\mathbf{r}, t) = \frac{i\hbar}{2m} (\Psi \nabla \Psi^* - \Psi^* \nabla \Psi) \quad (9.46)$$

The probability current density concept leads to the following important result: *If the probability of finding a particle in some region increases with time, the probability of finding it outside this region decreases by the same amount.*

9.16 NORMALIZATION CONDITION

Equation (9.45) which represents the sum of the probabilities over the whole space is called the normalization condition. This condition plays an important role in quantum mechanics, since it places a restriction on the possible solutions of the Schrödinger equation. In particular, the wave function $\Psi(\mathbf{r}, t)$ must approach zero sufficiently fast as \mathbf{r} tends to $\pm\infty$. Then only the integral in Eq. (9.45) remains finite.

In situations where $\Psi(\mathbf{r}, t)$ is not normalized, one can use this condition to normalize it. For that, multiply $\Psi(\mathbf{r}, t)$ by a constant N , called **normalization constant**, and substitute it in Eq. (9.45) as follows:

$$\int_{-\infty}^{\infty} (N^* \Psi^*) (N \Psi) d\tau = 1$$

$$|N|^2 \int_{-\infty}^{\infty} |\Psi|^2 d\tau = 1 \quad (9.47)$$

which is the normalization condition. Since the Schrödinger equation is linear, any solution multiplied by a constant is also a solution.

9.17 OPERATORS

An operator is defined as a rule by which a given function is transformed into another function. If the operator A transforms $f(x)$ into the function $g(x)$, we write it as

$$g(x) = A f(x) \quad (9.48)$$

In

$$g(x) = A f(x) = \frac{df(x)}{dx} \quad (9.49)$$

the operator A differentiates the function $f(x)$. In the equation

$$g(x) = A f(x) = [f(x)]^2 \quad (9.50)$$

the operator squares the function. If an operator A satisfies the following two properties, it is said to be a **linear operator**:

$$A[f_1(x) + f_2(x)] = A f_1(x) + A f_2(x) \quad (9.51)$$

$$A[c f(x)] = c A f(x) \quad (9.52)$$

where c is an arbitrary constant. All the operators occurring in quantum mechanics are linear.

There is another important type of operator known as **Hermitian operator**. Two important theorems regarding Hermitian operators which we use throughout quantum mechanics are:

- (i) The eigenvalues of Hermitian operators are real
- (ii) Eigenfunctions of a Hermitian operator that belong to different eigenvalues are orthogonal

9.18 HAMILTONIAN OPERATOR

The Hamiltonian of a particle is its total energy. We now derive the operator for the Hamiltonian of a particle of mass m moving in a potential $V(r)$. The kinetic energy T of the particle is given by

$$T = \frac{p^2}{2m} \quad (9.53)$$

The operator for p is $-i\hbar\nabla$. Replacing p by its operator in Eq. (9.53), we get

$$T_{op} = \frac{(-i\hbar\nabla) \cdot (-i\hbar\nabla)}{2m} = -\frac{\hbar^2}{2m} \nabla^2 \quad (9.54)$$

The operator for the potential energy $V(r)$ is $V(r)$ itself. Hence, the operator for the Hamiltonian H_{op} is given by

$$H_{op} = T_{op} + V_{op} = -\frac{\hbar^2}{2m} \nabla^2 + V(r) \quad (9.55)$$

For convenience, often the operator associated with dynamical variables will also be denoted by the same letter. Thus, the Hamiltonian operator

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V(r) \quad (9.56)$$

The operator for Hamiltonian is of fundamental importance in quantum mechanics.

9.19 EIGENVALUES AND EIGENFUNCTIONS OF AN OPERATOR

Often an operator A operating on a function $f_i(x)$ simply multiplies $f_i(x)$ by a constant a_i :

$$A f_i(x) = a_i f_i(x) \quad (9.57)$$

We then say that $f_i(x)$ is an eigenfunction of the operator A with *eigenvalue* a_i . In the equation

$$\frac{d}{dx} e^{5x} = 5e^{5x}$$

e^{5x} is an eigenfunction of the operator (d/dx) with the eigenvalue 5. In the equation

$$\frac{d^2}{d\phi^2} (\sin 5\phi) = -25 (\sin 5\phi)$$

$\sin 5\phi$ is an eigenfunction of the operator $(d^2/d\phi^2)$ with the eigenvalue -25 . The solution of such eigenvalue equations dominate quantum mechanics. If a given a_i is associated with a large number of f_i 's, the eigenvalue a_i is said to be **degenerate**.

9.19.1 Schrödinger Equation as an Eigenvalue Equation

An important eigenvalue equation is that of the Hamiltonian operator given in Eq. (9.55). Its eigenvalue equation corresponding to the eigenfunction $\psi(\mathbf{r})$ is given by

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (9.58)$$

which is simply the time independent Schrödinger equation in Eq. (9.43). The eigenvalue E in Eq. (9.58) is the energy of the system since the operator $\left(-\frac{\hbar^2}{2m} \right) \nabla^2 + V(r)$ represents the total energy. Hence, E in Eq. (9.43) which was introduced as a constant is simply the total energy of the system.

9.20 POSTULATES OF QUANTUM MECHANICS

In this section, we develop the foundation of quantum mechanics on the basis of certain postulates and then deduce their consequences. Although there is no proof for these postulates, one can validate the end results experimentally. There are many ways of stating the fundamental postulates of quantum mechanics. However, all are mathematically and physically equivalent.

Postulate 1: Wave function

*Every state of a dynamical system is described by a wave function Ψ of the co-ordinates and time. The function Ψ and its derivatives must be finite, single valued and continuous over the domain of the variables of Ψ . This function, called the **wave function** or **state function**, contains all the information that can be obtained about the system.*

The representation in which the wave function is a function of co-ordinates and time is called **co-ordinate representation**. We have already discussed the probability interpretation of Ψ in Section 9.15. If Ψ_1 and Ψ_2 are the wave functions of any two states of a system then their linear combination is given by

$$\Psi = c_1 \Psi_1 + c_2 \Psi_2, \quad c_1 \text{ and } c_2 \text{ are constants} \quad (9.58a)$$

which describes another state of the system. This result is known as the **superposition of states**.

Postulate 2: Operators

To every observable of classical mechanics, there corresponds a linear operator in quantum mechanics.

An observable is a physically measurable quantity. The operators associated with some of the observables are shown in Table 9.1. Operators representing energy, momentum, and potential energy have already been discussed in Section 9.13.

Table 9.1 Classical observables and their quantum mechanical operators

Classical observable	Quantum mechanical operator
Co-ordinates x, y, z	x, y, z
Function of co-ordinate $f(x, y, z)$	$f(x, y, z)$
Momentum p_x, p_y, p_z	$-i\hbar \frac{\partial}{\partial x}, -i\hbar \frac{\partial}{\partial y}, -i\hbar \frac{\partial}{\partial z}$
Momentum \mathbf{p}	$-i\hbar \nabla$
Energy E	$i\hbar \frac{\partial}{\partial t}$
Time t	t

Postulate 3: Expectation value

When a system is in a state represented by the normalized wave function Ψ , the **expectation value** of any observable a , whose operator is A , is given by

$$\langle a \rangle = \int_{-\infty}^{\infty} \Psi^* A \Psi d\tau \quad (9.59)$$

It may be noted here that the sandwiching of the operator between Ψ and Ψ^* is a necessity. If the wave function is not a normalized one, then

$$\langle a \rangle = \frac{\int_{-\infty}^{\infty} \Psi^* A \Psi dx}{\int_{-\infty}^{\infty} \Psi^* \Psi dx} \quad (9.60)$$

It can be proved that operators in quantum mechanics have real expectation values. If the wave function Ψ is an eigenfunction of the operator A corresponding to the eigenvalue α , then

$$A\Psi = \alpha\Psi \quad (9.61)$$

Therefore, Eq. (9.59) reduces to

$$\langle a \rangle = \int_{-\infty}^{\infty} \Psi^* \alpha \Psi dx = \alpha \int_{-\infty}^{\infty} \Psi^* \Psi dx = \alpha \quad (9.62)$$

That is, the expectation value is simply the eigenvalue. This leads us to the fourth postulate.

Postulate 4: Eigenvalues

The possible values which a measurement of an observable whose operator is A can give are the eigenvalues a_n of the equation

$$A\Psi_n = a_n \Psi_n, \quad n = 1, 2, 3, \dots \quad (9.63)$$

The eigenfunctions Ψ_n form a complete set of n independent functions.

A measurement always gives a real number. Hence, the eigenvalues have to be real as they are the ones we expect in a measurement. Since Hermitian operators have real eigenvalues, the operators associated with physical quantities must be Hermitian.

Postulate 5: Time development of a quantum system

The postulates considered so far describe a quantum system at a given instant of time. The fifth postulate describes the time development of the system (dynamics).

The wave function $\Psi(\mathbf{r}, t)$ of a system evolves in time according to the time dependent Schrödinger equation given by:

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = H \Psi(\mathbf{r}, t) \quad (9.64)$$

where H is the Hamiltonian operator which is independent of time.

This procedure of considering the state function $\Psi(\mathbf{r}, t)$ to depend on co-ordinates and time and the operator to be independent of time is called **Schrödinger representation**.

9.21 PARTICLE IN A ONE-DIMENSIONAL INFINITE SQUARE WELL

We shall now consider the solution of the time independent Schrödinger equation for a particle moving in a one dimensional square well of infinite depth as shown in Fig. 9.5(a). The potential $V(x)$ is given by

$$V(x) = \begin{cases} 0 & 0 \leq x \leq a \\ \infty & x < 0 \text{ and } x > a \end{cases} \quad (9.65)$$

which is the same for a one dimensional box. The Schrödinger equation, Eq. (9.43) for a one dimensional case takes the form

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi(x)}{dx^2} = (E - V) \Psi(x) \quad (9.65a)$$

In the regions $x < 0$ and $x > a$, the Schrödinger equation for the particle is:

$$\begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2\Psi}{dx^2} &= (E - \infty) \Psi(x) \equiv -\infty \Psi(x) \\ \Psi(x) &\underset{\infty}{=} \frac{1}{2m} \frac{d^2\Psi}{dx^2} = 0 \end{aligned} \quad (9.66)$$

That is, the wave function $\Psi(x) = 0$ outside the box.

9.21.1 Energy Eigenvalues and Eigenfunctions

Next, let us consider the region $0 \leq x \leq a$. The time independent Schrödinger equation for this region is:

$$\frac{-\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi \quad (9.67)$$

$$\frac{d^2\psi}{dx^2} = -k^2\psi; \quad k^2 = \frac{2mE}{\hbar^2} \quad (9.68)$$

The solution of this equation is

$$\psi = A \sin kx + B \cos kx \quad (9.69)$$

To determine the constants, we have to use the boundary conditions $\psi = 0$ at $x = 0$ and at $x = a$. The condition $\psi = 0$ at $x = 0$ leads to $B = 0$. Then,

$$\Psi = A \sin kx \quad (9.70)$$

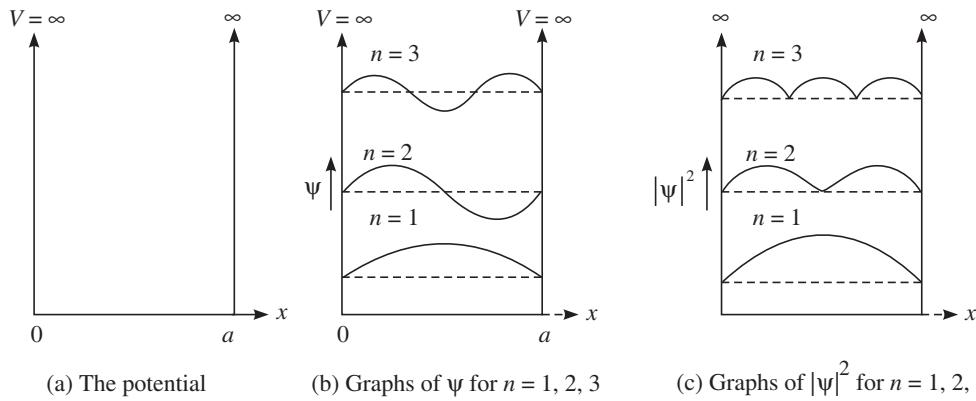


Fig. 9.5 A particle in a one dimensional box.

The condition $\psi = 0$ at $x = a$ gives $A \sin ka = 0$. The constant A cannot be zero, since this would make the function zero everywhere. Hence,

$$\sin ka = 0 \quad \text{or} \quad ka = n\pi, \quad n = 1, 2, 3, \dots \quad (9.71)$$

The value $n = 0$ is left out since it leads to $\psi = 0$. From Eqs. (9.71) and (9.68), we have

$$\begin{aligned} k &= \frac{n\pi}{a} \quad \text{or} \quad \frac{2mE_n}{\hbar^2} = \frac{n^2\pi^2}{a^2} \\ E_n &= \frac{n^2\pi^2\hbar^2}{2ma^2}, \quad n = 1, 2, 3, \dots \end{aligned} \quad (9.72)$$

These are the energy eigenvalues for the particle in the one dimensional box.

With the above value of k , the wave function ψ , Eq. (9.70), becomes

$$\psi = A \sin \frac{n\pi x}{a}, \quad n = 1, 2, 3, \dots \quad (9.73)$$

Use of the normalization condition gives

$$\begin{aligned} |A|^2 \int_0^a \sin^2 \frac{n\pi x}{a} dx &= 1 \\ |A|^2 \cdot \frac{a}{2} &= 1 \quad \text{or} \quad A = \sqrt{\frac{2}{a}} \end{aligned} \quad (9.74)$$

The eigenfunction of the different states of the particle in the one dimensional box is given as

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}, \quad n = 1, 2, 3, \dots \quad (9.75)$$

The plots of $\psi(x)$ against x for $n = 1, 2$ and 3 are illustrated in Fig. 9.5(b).

9.21.2 Probability Distribution

$$\text{The probability density } P(x) = |\psi|^2 = \frac{2}{a} \sin^2 \frac{n\pi x}{a} \quad (9.76)$$

The plot of probability density $|\psi|^2$ against x for $n = 1, 2$ and 3 are illustrated in Fig. 9.5(c). From Eq. (9.72), we get

$$E_1 = \pi^2 \hbar^2 / (2ma^2).$$

It may be noted that the lowest or ground state energy is not zero as expected classically. This energy is called the **zero point energy**. It is a consequence of the uncertainty principle.

9.22 QUANTUM MECHANICAL TUNNELLING

Quantum mechanical tunnelling of barrier was first used by Gamow and others in 1928 to explain α -decay by nuclei. To understand the phenomenon of tunnelling, we shall consider a typical example. The barrier potential, illustrated in Fig. 9.6(a), is given by

$$V(x) = \begin{cases} V_0 & 0 \leq x \leq a \\ 0 & x < 0 \quad \text{and} \quad x > a \end{cases} \quad (9.77)$$

Consider particles of mass m and energy $E < V_0$ incident on the barrier from left. Classically, all the particles will be reflected at $x = 0$, if $E < V_0$ and all will be transmitted into the region $x > a$ if $E > V_0$. However, quantum mechanically there will be a finite probability for the particles to be in region 3 even if $E < V_0$. This phenomenon of tunnelling through the barrier is called **barrier penetration** or **quantum mechanical tunnelling**. This tunnelling through the barrier is possible because of the wave nature of material particles. Though the particle tunnels through the region $0 < x < a$, it can never be observed there as its kinetic energy is negative in that region.

The solution of the Schrödinger equation for the particle gives the wave function for the three regions. In region 1, the wave function corresponds to two waves, one travelling from

left to right and the other from right to left. These represent the incident particles and the particles reflected by the barrier, respectively. In region 2, two solutions are possible—an exponentially increasing function and an exponentially decreasing function. The wave function in region 3 is finite, and we have only waves travelling from left to right. Figure 9.6(b) illustrates the form of the wave function in the three regions. Thus, quantum mechanically there is a probability for the particle to be in region 3.

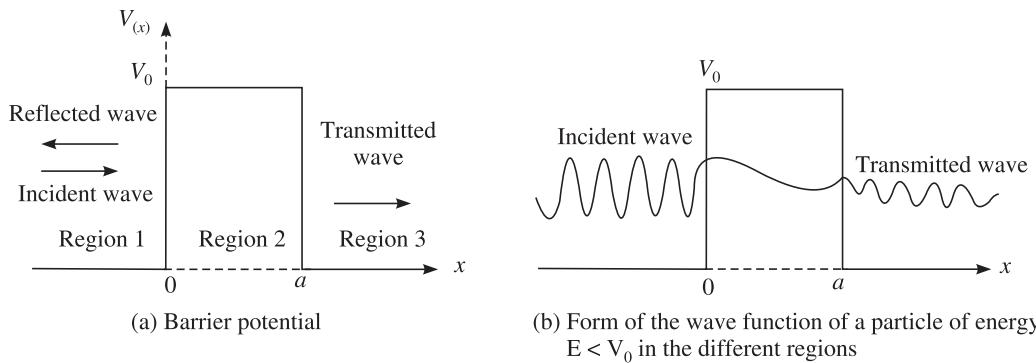


Fig. 9.6 Quantum mechanical tunnelling.

The exponentially increasing solution in region 2 and the reflected wave in region 1 are not shown.

The concept of barrier penetration is used to explain a number of phenomena in physics. Some of them are as follows:

- (i) The emission of α -particles from radioactive nuclei.
- (ii) Periodic inversion in ammonia molecule which is the basis for the atomic clock.
- (iii) Barrier penetration in electronic devices such as tunnel diode and Josephson junction.
- (iv) Electron tunnelling in scanning tunnelling microscope.

SOLVED EXAMPLES

Example 9.1 Determine the wavelength of an electron accelerated from rest through a potential difference of 100 volts.

Solution: The de Broglie wavelength λ is given by

$$\begin{aligned}\lambda &= \sqrt{\frac{150}{V}} \times 10^{-10} \text{ m} \\ &= \sqrt{\frac{150}{100}} \times 10^{-10} \text{ m} = 1.22 \times 10^{-10} \text{ m}\end{aligned}$$

Example 9.2 Compute the de Broglie wavelength of an electron whose kinetic energy is 10 eV.

Solution: The kinetic energy $E_k = \frac{p^2}{2m}$

$$\begin{aligned} p &= \sqrt{2m E_k} = \sqrt{2(9.11 \times 10^{-31} \text{ kg})(10 \text{ eV})(1.6 \times 10^{-19} \text{ J/eV})} \\ &= 1.707 \times 10^{-24} \text{ kg m/s} \end{aligned}$$

$$\lambda = \frac{h}{p} = \frac{6.626 \times 10^{-34} \text{ Js}}{1.707 \times 10^{-24} \text{ kg m/s}} = 3.88 \times 10^{-10} \text{ m} = 0.39 \text{ nm}$$

Example 9.3 An electron and a proton have the same non-relativistic kinetic energy. Show that the proton has a shorter de Broglie wavelength.

Solution: Kinetic energy $E_k = \frac{p^2}{2m}$ or $p = \sqrt{2m E_k}$

$$\text{The de Broglie wavelength } \lambda = \frac{h}{p} = \frac{h}{\sqrt{2m E_k}}$$

$$\text{Since both have the same kinetic energy, } \frac{\lambda_e}{\lambda_p} = \sqrt{\frac{m_p}{m_e}}$$

Since $m_p > m_e$, we have $\lambda_p < \lambda_e$

Example 9.4 An electron moves with a constant speed $v = 1.1 \times 10^6 \text{ m/s}$ in a straight line. If the speed is measured to a precision of 0.1 per cent, what is the maximum precision with which its position could be simultaneously measured?

Solution: The momentum of the electron $p = mv$

$$p = (9.11 \times 10^{-31} \text{ kg})(1.1 \times 10^6 \text{ m/s}) = 10^{-24} \text{ kg m/s}$$

$$\Delta p = \frac{0.1}{100} \times 10^{-24} \text{ kg m/s} = 10^{-27} \text{ kg m/s}$$

From Eq. (9.16), we get

$$\Delta x = \frac{\hbar}{2\Delta p} = \frac{1.054 \times 10^{-34} \text{ Js}}{2 \times 10^{-27} \text{ kg m/s}} = 0.53 \times 10^{-7} \text{ m}$$

Example 9.5 An electron remains in an excited state of an atom for 10^{-8} s . What is the minimum uncertainty in the energy of the state in eV?

Solution: We have $(\Delta E)(\Delta t) \geq \frac{\hbar}{2}$; $\Delta t \sim 10^{-8} \text{ s}$

$$\begin{aligned}\Delta E &\geq \frac{1.05 \times 10^{-34} \text{ Js}}{2(10^{-8} \text{ s})} = \frac{1.05}{2} \times 10^{-26} \text{ J} \\ &= \frac{1.05 \times 10^{-26} \text{ J}}{2(1.6 \times 10^{-19} \text{ J/eV})} = 3.28 \times 10^{-7} \text{ eV}\end{aligned}$$

Example 9.6 The average lifetime of an excited atomic state is 10^{-8} s . If the wavelength of the spectral line associated with a transition from this state is 400 nm , what is the width of this line?

Solution: Average lifetime is 10^{-8} s . Hence, $\Delta t \sim 10^{-8} \text{ s}$

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

Also, we have

$$\Delta E \equiv \frac{\hbar}{2\Delta t}$$

Equating the two values of ΔE , we get $\frac{hc}{\lambda^2} \Delta\lambda = \frac{\hbar}{2 \cdot \Delta t}$

$$\text{or } \Delta\lambda = \frac{\lambda^2}{4\pi c \Delta t} = \frac{(400 \times 10^{-9} \text{ m})^2}{4\pi (3 \times 10^8 \text{ m/s}) (10^{-8} \text{ s})} = 4.24 \times 10^{-15} \text{ m}$$

Example 9.7 The uncertainty in the velocity of a particle is equal to its velocity. If $(\Delta x)(\Delta p) \equiv h$, show that the uncertainty in its location is its de Broglie wavelength.

Solution: It is given that $\Delta v = v$. Therefore, we have

$$p = mv \quad \text{or} \quad \Delta p = m \Delta v = mv$$

Substituting this value of Δp in the given relation, we get

$$(\Delta x)mv = h \quad \text{or} \quad \Delta x = \frac{h}{mv}$$

which is its de Broglie wavelength.

Example 9.8 A quantum mechanical system is described by the wave function

$$\psi(x) = \left(2/a_0^{3/2}\right) xe^{-x/a_0}; \quad a_0 \text{ constant}; \quad 0 \leq x \leq \infty$$

Calculate the probability density and show that the wave function is normalized. It is given that

$$\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}}.$$

Solution: The probability density $P(x) = \psi\psi^* = |\psi|^2 = \frac{4}{a_0^3} x^2 e^{-\frac{2x}{a_0}}$

$$\begin{aligned} \text{The probability integral } \int_0^\infty |\psi|^2 dx &= \frac{4}{a_0^3} \int_0^\infty x^2 e^{-\frac{2x}{a_0}} dx \\ &= \frac{4}{a_0^3} \frac{2!}{\left(\frac{2}{a_0}\right)^3} = 1 \end{aligned}$$

which is the normalization condition. Hence, the given wave function is normalized.

Example 9.9 A particle constrained to move along x -axis in the region $0 \leq x \leq a$ has a wave function $\Psi(x) = N \sin\left(\frac{n\pi x}{a}\right)$, where n is an integer. Normalize the wave function.

Solution: The normalization condition gives

$$\begin{aligned} N^2 \int_0^a \sin^2\left(\frac{n\pi x}{a}\right) dx &= 1 \\ N^2 \int_0^a \frac{1}{2} \left(1 - \cos\frac{2n\pi x}{a}\right) dx &= 1 \\ N^2 \frac{a}{2} &= 1 \quad \text{or} \quad N = \sqrt{\frac{2}{a}} \end{aligned}$$

$$\text{The normalized wave function } \psi(x) = \sqrt{\frac{2}{a}} \sin\frac{n\pi x}{a}$$

Example 9.10 If the wave function $\psi(x) = A \sin kx$ satisfies the time independent Schrödinger equation, find the form of the potential $V(x)$.

Solution: The wave function $\psi(x) = A \sin kx$ satisfies the Schrödinger equation.

This implies

$$\begin{aligned} \left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] A \sin kx &= EA \sin kx \\ \left[\frac{\hbar^2 k^2}{2m} + V(x) \right] A \sin kx &= EA \sin kx \\ \frac{\hbar^2 k^2}{2m} + V(x) &= E \quad \text{or} \quad V(x) = E - \frac{\hbar^2 k^2}{2m} \end{aligned}$$

Example 9.11 The functions $\psi_1(x)$ and $\psi_2(x)$ are the normalized eigenfunctions of an operator A having the same eigenvalue a . Show that the function $\psi = c_1 \psi_1 + c_2 \psi_2$, c_1 and c_2 are constants, is also an eigenfunction of A with the same eigenvalue.

Solution: We have the eigenvalue equations given as

$$A\Psi_1 = a\Psi_1 \quad \text{and} \quad A\Psi_2 = a\Psi_2$$

Consider the eigenvalue equation with the function $\Psi = c_1\Psi_1 + c_2\Psi_2$

$$\begin{aligned} A(c_1\Psi_1 + c_2\Psi_2) &= c_1A\Psi_1 + c_2A\Psi_2 = c_1a\Psi_1 + c_2a\Psi_2 \\ &= a(c_1\Psi_1 + c_2\Psi_2) \end{aligned}$$

Hence, the result.

Example 9.12 Show that the expectation value $\langle p_x \rangle$ is real

Solution: We have

$$\begin{aligned} \langle p_x \rangle &= \int_{-\infty}^{\infty} \Psi^* \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi dx = -i\hbar \int_{-\infty}^{\infty} \Psi^* \frac{\partial \Psi}{\partial x} dx \\ &= -i\hbar \int_{-\infty}^{\infty} \left(\Psi^* \frac{\partial \Psi}{\partial x} + \frac{\partial \Psi^*}{\partial x} \Psi \right) dx + i\hbar \int_{-\infty}^{\infty} \frac{\partial \Psi^*}{\partial x} \Psi dx \end{aligned}$$

where we have added the second term and subtracted the same term in the first term.

$$\begin{aligned} \langle p_x \rangle &= -i\hbar \int_{-\infty}^{\infty} \frac{\partial}{\partial x} (\Psi^* \Psi) dx + \int_{-\infty}^{\infty} \Psi \left(i\hbar \frac{\partial}{\partial x} \right) \Psi^* dx \\ &= -i\hbar \left[\Psi^* \Psi \right]_{-\infty}^{\infty} + \langle p_x \rangle^* = \langle p_x \rangle^* \end{aligned}$$

since the wave function vanishes at infinity the first term is zero. Hence, $\langle p_x \rangle$ is real.

Example 9.13 Calculate the expectation value of the co-ordinate x corresponding to the wave

$$\text{function } \Psi_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}, \quad 0 \leq x \leq a, \quad n = 1, 2, \dots$$

$$\begin{aligned} \textbf{Solution: } \langle x \rangle &= \frac{2}{a} \int_0^a x \sin^2 \frac{n\pi x}{a} dx = \frac{2}{a} \int_0^a x \left(1 - \cos \frac{2n\pi x}{a} \right) dx \\ &= \frac{1}{a} \int_0^a x dx - \frac{1}{a} \int_0^a x \cos \frac{2n\pi x}{a} dx \end{aligned}$$

The second term is zero since

$$\int_0^a x \cos \frac{2n\pi x}{a} dx = \frac{a}{2n\pi} \left[x \sin \frac{2n\pi x}{a} \right]_0^a - \frac{a}{2n\pi} \int_0^a \sin \frac{2n\pi x}{a} dx$$

$$= 0 + \left(\frac{a}{2n\pi} \right)^2 \left[\cos \frac{2n\pi x}{a} \right]_0^a = 0$$

$$\langle x \rangle = \frac{1}{a} \cdot \frac{a^2}{2} = \frac{a}{2}$$

Example 9.14 An electron is moving in a one dimensional box of width 2×10^{-10} m. If the wave function of the first excited state $\psi(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{2\pi x}{a}\right)$, what is the probability of finding the electron between $x = 0$ and $x = 10^{-10}$ m in that state.

Solution: The required probability $= \frac{2}{a} \int_{x_1}^{x_2} \sin^2\left(\frac{2\pi x}{a}\right) dx$

where $x_1 = 0$, $x_2 = 1 \times 10^{-10}$ m, $a = 2 \times 10^{-10}$ m

$$\begin{aligned} \text{Probability} &= \frac{2}{a} \int_{x_1}^{x_2} \frac{1}{2} \left(1 - \cos \frac{4\pi x}{a} \right) dx \\ &= \frac{2}{a} \cdot \frac{1}{2} x_2 - \frac{2}{a} \cdot \frac{1}{2} \cdot \frac{a}{4\pi} \sin \frac{4\pi x_2}{a} = \frac{1}{2} \end{aligned}$$

REVIEW QUESTIONS

- 9.1 Why do we say that light has both wave and particle properties?
- 9.2 Explain de Broglie's hypothesis.
- 9.3 What are matter waves?
- 9.4 If an electron and a proton travel at the same speed, which has the shorter de Broglie wavelength?
- 9.5 If an electron and a proton have the same kinetic energy, which has the shorter de Broglie wavelength?
- 9.6 Why are both particle and wave aspects of material particles needed for a complete description of physical phenomena?
- 9.7 Does the concept of Bohr orbit violate the uncertainty principle?
- 9.8 What is a wave packet? How is it represented geometrically?
- 9.9 State and explain the de Broglie relation.
- 9.10 State and explain the uncertainty principle.
- 9.11 Why does the frequency of light emitted by an atom have a width?

- 9.12** It is said that the ground state energy of hydrogen atom can be known precisely, but the excited states have some uncertainty in their values. Is this consistent with the uncertainty principle? Explain.
- 9.13** Outline the probability interpretation of the wave function.
- 9.14** What is the physical significance of the integral $\int_{-\infty}^{\infty} |\psi|^2 dx = 1$?
- 9.15** Why is it necessary for a wave function to be normalized?
- 9.16** What do you understand by co-ordinate representation?
- 9.17** Obtain the operator for the Hamiltonian of a particle moving in a potential $V(r)$.
- 9.18** What are the eigenvalues and eigenfunctions of an operator? What is the significance of the eigenvalue?
- 9.19** Explain the postulate that describes the time development of a system.
- 9.20** Define the expectation value of a dynamical variable in quantum mechanics.
- 9.21** Write a note on quantum mechanical tunnelling.
- 9.22** What do you understand by the wave function of a particle? Is it possible to determine it? Explain.
- 9.23** What are matter waves? Obtain an expression for de Broglie wavelength. Derive expressions for the de Broglie wavelength of an electron (i) accelerated from rest through a potential of V volts (ii) having kinetic energy T .
- 9.24** Explain Heisenberg's uncertainty principle. How does the uncertainty principle account for the absence of electrons in the nucleus?
- 9.25** Write down the Schrödinger equation for a particle in a one-dimensional box. Solve the same to obtain its energy eigenvalues and probability density.
- 9.26** State and explain the different postulates of quantum mechanics.
- 9.27** Derive the time dependent Schrödinger equation for a particle moving in a potential $V(r)$. Hence derive the time independent Schrödinger equation.
- 9.28** Assuming the time dependent Schrödinger equation, derive the time independent Schrödinger equation for a particle moving in a potential $V(r)$. Explain the physical significance of the wave function.

PROBLEMS

- Electrons in an electron microscope are accelerated from rest through a potential difference of V volts so that their de Broglie wavelength is 0.04 nm. What is the value of V ?
- Compute the de Broglie wavelength of (i) an electron of kinetic energy 4.5 keV and (ii) a proton of kinetic energy 4.5 keV.

3. What voltage is needed to produce electron wavelength of 0.10 nm?
4. Calculate the ratio of the kinetic energy of an electron to that of a proton if their de Broglie wavelengths are equal.
5. The position and momentum of a 2 keV electron are determined simultaneously. If its position is located within 1 Å, what is the percentage of uncertainty in its momentum?
6. Wavelengths can be determined with an accuracy of one in 10^7 . What is the uncertainty in the position of a 2 Å photon when its wavelength is simultaneously measured?
7. If an electron's position can be measured to an accuracy of 2.0×10^{-8} m, how accurately can its velocity be known?
8. A quantum mechanical system having mass m is described by the wave function

$$\Psi(x) = \left(\frac{m\omega}{\pi\hbar} \right)^{\frac{1}{4}} \exp\left(\frac{-m\omega x^2}{2\hbar} \right), \quad -\infty \leq x \leq \infty$$

Calculate the probability density and show that the wave function is normalized. Given

that $\int_0^\infty e^{-\alpha x^2} dx = \left(\frac{\pi}{4\alpha} \right)^{\frac{1}{2}}$.

9. A particle constrained to move along x -axis in the domain $0 \leq x \leq a$ has a wave function $\Psi(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right)$. Evaluate the expectation value of x and the x -component of its momentum.
10. For an electron in a one dimensional box of width 1 Å, calculate the separation between the two lowest energy levels in electron volt.
11. Obtain the operator for the x , y and z components of the orbital angular momentum.
12. Is $\Psi(x, t) = A \sin(kx - \omega t)$ a solution of the time-dependent Schrödinger equation?

C H A P T E R

10

Statistical Mechanics

In physics, it is the practice to take a number of measurements of an observable and then to evaluate its mean value or the most probable value. Statistical mechanics makes no attempt to obtain a complete solution for complex systems. Instead, it makes predictions about the measurable properties of macroscopic systems by studying the probable behaviour of the system from statistical considerations. This approach to explain the behaviour of large systems is called **statistical mechanics**.

10.1 STATISTICAL METHODS

10.1.1 Macroscopic and Microscopic Systems

The physical world that we experience consists of macroscopic objects which are systems that are large compared to atomic dimensions. Let us consider a gas confined in a container. It is the collection of a very large number of molecules or atoms, perhaps of the order of 10^{20} or more. A gas molecule or atom in a container is a microscopic system, whereas the gas in the container is a macroscopic system. One way to analyse such a complex system is to specify a set of microscopic properties such as the position and velocity of each gas molecule. When the number of particles is large as in this case, such an analysis is highly impractical. However, we can predict the values of the measurable properties of macroscopic systems without tracking the motions of each individual particle. This shortcut is possible as we can apply general principles of physics such as conservation of energy, conservation of momentum, etc., to a large collection of particles. Then, relations connecting microscopic and macroscopic properties help to understand the microscopic properties of the system. In the case of a gas confined in a container, kinetic theory gives the relationship between the microscopic motion of the molecules and the macroscopic temperature and pressure.

10.1.2 Macrostate and Microstate

Generally, one can make a statistical analysis by counting the number of different arrangements of the microscopic properties of a system. As an example, consider the distribution of 2 units of energy to a gas of 4 identical distinguishable particles. Let us impose a condition that each particle can acquire energy only in integral units. One way for the particles to share the energy is to take the entire two units by one particle. There are four ways of having this distribution, as shown in Table 10.1 Another way to distribute the energy is to give two different particles 1 unit each. This distribution can be obtained in six ways (Table 10.1). Each possible energy distribution is called a **macrostate**. In the example considered there are 2 macrostates, one in which one particle has two units of energy and another in which 2 particles have 1 unit of energy. The different arrangements of microscopic variables corresponding to a single macrostate are called **microstates**. In the example considered, there are four microstates corresponding to macrostate A and six microstates corresponding to macrostate B. The number of microstates corresponding to a given macrostate is called the **multiplicity**.

Table 10.1 Macro and microstates of a simple system

Macrostate	Microstate			
	Energy of particle			
	1	2	3	4
A	2	0	0	0
	0	2	0	0
	0	0	2	0
	0	0	0	2
B	1	1	0	0
	1	0	1	0
	1	0	0	1
	0	1	1	0
	0	1	0	1
	0	0	1	1

10.1.3 Phase Space

According to classical mechanics, at any time the state of a system may be completely defined by specifying the position and velocities of all its component parts. For a system having f degrees of freedom the state can be specified by means of f generalised co-ordinates q_1, q_2, \dots, q_f , and f generalised momenta p_1, p_2, \dots, p_f . The conceptual space having $2f$ dimensions formed by these f co-ordinates and f generalised momenta is called the **phase space**. The instantaneous state of a system having f degrees of freedom can be specified by a point called **phase point** in this $2f$ dimensional space. The volume element $d\tau$ in phase space is given by $d\tau = dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f$.

10.1.4 Cells in Phase Space

To apply classical statistics to a system of molecules, the phase space available for a molecule having f degrees of freedom is divided into a number of elementary regions or cells of volume

$$dq_1 dq_2 dq_3 \dots dq_f dp_1 dp_2 dp_3 \dots dp_f$$

in the $2f$ dimensional space. All the cells have equal extensions, but their position in the space will differ corresponding to the different values of the co-ordinates $q_1, q_2 \dots q_f, p_1, p_2 \dots p_f$ of the molecule. Each molecule may lie in a different cell if its co-ordinates and momenta are different. It is also possible for several molecules to occupy the same cell if their co-ordinates and momenta happen to fall within the same range. Each arrangement of specified individual molecule having its representative point in particular cells is a microscopic state of the system. The macroscopic state of n molecules is determined by specifying the number of molecules whose representative points are to be found in the different unit cells. Based on cells in phase space, the theory of classical statistics was developed.

10.2 CLASSICAL STATISTICS

Statistical physics is mainly concerned with the distribution of a fixed quantity of energy among a large number of particles, from which the observable properties of the system may be deduced. Let us consider a system having a large number of identical, but distinguishable particles in thermal equilibrium. Boltzmann derived a distribution relation that predicts the probable numbers of identical particles that will occupy each of the available energy cells or states. The Boltzmann distribution function $f_B(E)$ is given by

$$f_B(E) = A e^{-E/kT} \quad (10.1)$$

where A is a constant, E is the energy of the state, and k is the Boltzmann constant.

$$k = 1.38 \times 10^{-23} \text{ J/K} = 8.62 \times 10^{-5} \text{ eV/K}$$

Similar results were obtained earlier by Maxwell for speed distribution. The factor $\exp\left(-\frac{E}{kT}\right)$ is often referred to as the Boltzmann factor. The degeneracy of the state with energy E , often referred to as the density of states, is usually denoted by $g(E)$. Then $g(E) dE$ gives the number of states with energy between E and $E + dE$. The number of particles with energy E is then given by

$$N(E) = g(E)f_B(E) = Ag(E) e^{-\frac{E}{kT}} \quad (10.2)$$

The Boltzmann distribution is a very fundamental relation that can be used to explain properties of classical systems belonging to gases and condensed matter. The equipartition theorem which states that in equilibrium, each degree of freedom contributes $\frac{1}{2} kT$ to the average energy per molecule is derivable from the Boltzmann distribution. However, phenomena such as electrical conductivity, specific heat of metals, the behaviour of liquid helium and thermal radiation could not be successfully analysed. For these phenomena one has to use the methods of quantum statistics.

10.3 QUANTUM STATISTICS

In classical statistics, the molecules of a particular type, although similar, are distinguishable from one another. This implies that the process of molecular interchange gives a new microscopic state. However, according to quantum mechanics, particles are described by wave functions. The finite extent and overlap of wave functions make identical particles indistinguishable. It is this property of indistinguishability that makes quantum statistics different from classical statistics.

The symmetry considerations of the wave functions restrict them into two types: a symmetric wave function in the exchange of particles, and an antisymmetric wave function. Hence, the indistinguishable particles are divided into two types. Particles with half integral spin are described by an antisymmetric wave function, and those with integral spin (including zero) are described by a symmetric wave function. Systems with integral spin particles obey Bose-Einstein statistics, which was developed by S.N. Bose and Einstein. These particles are referred to as **bosons**. Systems with half integral particles obey Fermi-Dirac statistics. Hence, half integral particles are called **fermions**. This statistics was developed by Fermi and Dirac. Electrons, protons, and neutrons (all spin 1/2 particles) are fermions, while photons (spin 1), helium atom (spin 0), and π mesons (spin 0) are bosons. The symmetry restriction on the wave function are not there if the particles are non-interacting and distinguishable. This statistics is referred to as **Maxwell-Boltzmann statistics**.

10.4 BASIC POSTULATES OF QUANTUM STATISTICS

In quantum mechanics, the different wave functions of a system represent different probability densities. That is, each wave function corresponds to an eigenstate of the system. The eigenstate of quantum mechanics is the equivalent of the unit cell in classical statistics. Although Maxwell Boltzmann is a classical distribution, here it is considered with quantum considerations. We now sum up the postulates of Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac statistics as follows:

Maxwell—Boltzmann statistics

- (i) Identical particles of systems are distinguishable.
- (ii) A state can accommodate any number of particles.
- (iii) Applicable to both integral and half integral spin systems.
- (iv) Particles are non-interacting (ideal case) or weakly interacting.

Bose—Einstein statistics

- (i) Identical particles of systems are indistinguishable.
- (ii) Applicable to systems having integral spin.
- (iii) Applicable to identical particles whose wave function is symmetric in the exchange of particles.
- (iv) A quantum state can accommodate any number of bosons. That is, the particles need not obey Pauli's exclusion principle.
- (v) Particles are non-interacting (ideal case) or weakly interacting.

Fermi—Dirac statistics

- (i) Identical particles of systems are indistinguishable.
- (ii) Applicable to systems having half integral spin.
- (iii) Applicable to identical particles whose wave function is anti-symmetric in the exchange of particles.
- (iv) A quantum state can accommodate only one particle. If spin includes 2 particles, Pauli's principle is obeyed.
- (v) Particles are non-interacting (ideal case) or weakly interacting.

10.5 DISTRIBUTION LAWS OF THE DIFFERENT STATISTICS

In all the three cases, a system consisting of n identical particles may be considered. The total energy E of the system is constant. The n particles are to be distributed so that n_1 particles are in state 1, n_2 in state 2, n_3 in state 3, and so on. The degeneracy of the i th level is g_i . The energy of state 1 is E_1 , that of state 2 is E_2 , that of state 3 is E_3 , and so on.

10.5.1 Maxwell-Boltzmann Statistics

In Maxwell—Boltzmann, although the particles are identical they are distinguishable. As mentioned, there are no symmetry restrictions on the wave functions and therefore a state can accommodate any number of particles. Based on the assumptions mentioned in Section 10.4 and using Stirling's approximation and Lagrange method of undetermined multipliers, one can derive the distribution law given as

$$n_i = \frac{g_i}{e^{(\alpha + \beta E_i)}} \quad (10.3)$$

where e^α is a system dependent constant and $\beta = \frac{1}{kT}$, k being Boltzmann constant.

Often, the number of particles $n(E)$ with energy E is needed. The quantity $n(E)dE$ is then the number of particles in the energy range E and $E + dE$. In such situations, the distribution law, Eq. (10.3), can be written as

$$n(E) = g(E) f_{MB}(E) \quad (10.4)$$

where the function $g(E)$, called **density of states**, is the number of states available per unit energy range. The function $n(E)/g(E)$ is the **distribution function**. A comparison of Eqs. (10.3) and (10.4) and replacement of the constant e^α by A gives

$$f_{MB}(E) = \frac{1}{A e^{\frac{E}{kT}}} \quad (10.5)$$

which is referred to as the **Maxwell—Boltzmann distribution function**.

10.5.2 Bose-Einstein Statistics

In Bose—Einstein statistics, the n identical particles are indistinguishable. Since the particles are of integral spin, a state can accommodate any number of particles. The theoretical treatment

for photons done by Bose, in 1924, was generalised by Einstein which resulted in the distribution law given as

$$n_i = \frac{g_i}{e^{(\alpha+\beta E_i)} - 1} \quad (10.6)$$

Again e^α is a system dependent constant and $\beta = \frac{1}{kT}$. The Bose-Einstein distribution function $f_{BE}(E)$ becomes

$$f_{BE}(E) = \frac{1}{B_1 e^{\frac{E}{kT}} - 1} \quad (10.7)$$

where we have replaced $\exp(\alpha)$ by B_1 .

10.5.3 Fermi-Dirac Statistics

Similar to Bose-Einstein statistics, in the Fermi-Dirac statistics the n particles constituting the system are indistinguishable. Following Dirac's development of relativistic quantum mechanics, Fermi and Dirac completed the statistical mechanics of indistinguishable particles obeying Pauli's exclusion principle. The result is the Fermi-Dirac distribution law.

$$n_i = \frac{g_i}{e^{(\alpha+\beta E_i)} + 1} \quad (10.8)$$

where $\beta = \frac{1}{kT}$. Replacing the system dependent $\exp(\alpha)$ by B_2 , the Fermi-Dirac distribution function $f_{FD}(E)$ becomes

$$f_{FD}(E) = \frac{1}{B_2 e^{\frac{E}{kT}} + 1} \quad (10.9)$$

which is obeyed by large ensembles of identical indistinguishable particles obeying the exclusion principle. At high energies, both $f_{BE}(E)$ and $f_{FD}(E)$ approach $f_{MB}(E)$.

10.6 COMPARISON OF THE THREE STATISTICS

For comparison, the distribution law for the three statistics can be written in the following forms.

Bose-Einstein statistics

$$\begin{aligned} n_i &= \frac{g_i}{e^{\alpha+\beta E_i} - 1} \\ \frac{g_i}{n_i} + 1 &= e^\alpha e^{\frac{E_i}{kT}} \end{aligned} \quad (10.10)$$

Fermi-Dirac statistics

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

$$\frac{g_i}{n_i} - 1 = e^\alpha e^{\frac{E_i}{kT}} \quad (10.11)$$

Maxwell-Boltzmann statistics

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

$$\frac{g_i}{n_i} = e^\alpha e^{\frac{E_i}{kT}} \quad (10.12)$$

It is evident from these three equations that all the three statistics are likely to give the same result if $\left(\frac{g_i}{n_i}\right) = \frac{g(E)}{n(E)}$ is very large in comparison with unity. In such a case

$$\frac{g_i}{n_i} \pm 1 \approx \frac{g_i}{n_i} \quad (10.13)$$

When this condition is satisfied the expressions for distribution law in the three statistics will become equal. Under normal conditions this requirement is satisfied by most of the gases and liquids and the classical Maxwell-Boltzmann statistics will be adequate. There are a limited number of cases in which this distribution is not applicable, the radiation, liquid helium II and electron gas in metals. At low temperatures and at high pressures, the Maxwell-Boltzmann statistics fails for almost all systems.

10.7 DENSITY OF STATES

As mentioned in Section 10.5, the density of states $g(E)$ is the number of states available per unit energy range. To derive an expression for $g(E)$ the relationship between the number of unit cells and the number of eigenstates would be useful.

10.7.1 Phase Volume and Eigenstates

The volume element for a system having f degrees of freedom in the phase space is $dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f$. Hence,

$$\text{The total phase volume} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f \quad (10.14)$$

The dimension of $pq = \text{ML}^2 \text{T}^{-1}$. Therefore, for a system having f degrees of freedom, the dimension of volume in phase space = $[\text{ML}^2 \text{T}^{-1}]^f$. The dimension of Planck's constant is also

$\text{ML}^2 \text{T}^{-1}$. Hence, it is reasonable to identify the volume of a unit cell in phase space with the quantity h^f .

$$\text{Volume of a unit cell} = h^f \quad (10.15)$$

The number of unit cells in the phase space which is taken as the number of eigenstates g for a system is given as

$$g = \frac{1}{h^f} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f \quad (10.16)$$

A number of deductions based on this result were proved experimentally, which justifies this result.

10.7.2 Number of Eigenstates in an Energy Range

Often, the number of eigenstates in the energy range E and $E + dE$, where dE is small, but large enough to contain an appreciable number of energy levels, is needed. This can be evaluated using the result in Eq. (10.16). The volume of a unit cell in phase space is h^f , where f is the degrees of freedom of the particle. A point particle in space has three degrees of freedom. Let us denote the three co-ordinates by x , y , and z and the corresponding momenta by p_x , p_y , and p_z . The number of eigenstates

$$g = \frac{1}{h^3} \iiint dx dy dz \iiint dp_x dp_y dp_z \quad (10.17)$$

If V is the volume of the containing vessel, then

$$g = \frac{V}{h^3} \iiint dp_x dp_y dp_z \quad (10.18)$$

We are interested in the energy range E and $E+dE$ which corresponds to momentum range p and $p+dp$. Hence, $\iiint dp_x dp_y dp_z$ is simply the volume enclosed between the spheres of radii p and $p + dp$ given as:

$$\iiint dp_x dp_y dp_z = \frac{4}{3} \pi(p + dp)^3 - \frac{4}{3} \pi p^3 \equiv 4\pi p^2 dp \quad (10.19)$$

since dp is small. We have

$$E = \frac{p^2}{2m} \quad \text{or} \quad p = \sqrt{2mE} \quad (10.20)$$

$$dE = \frac{pdः}{m} \quad \text{or} \quad pdः = mdE$$

$$p^2 dp = pdः = m\sqrt{2mE} dE \quad (10.21)$$

Since we have evaluated the integral in the energy range E and $E + dE$, the number of states we get correspond to the energy range E and $E + dE$. From Eqs. (10.18), (10.19), and (10.21), the number of eigenstates in the energy range E and $E + dE$ is given by

$$\begin{aligned}
g(E) dE &= \frac{V}{h^3} 4 \pi m \sqrt{2mE} dE \\
&= \frac{4\pi m V}{h^3} (2mE)^{\frac{1}{2}} dE
\end{aligned} \tag{10.22}$$

The density of states is given by

$$g(E) = \frac{4\pi m V}{h^3} (2mE)^{\frac{1}{2}} \tag{10.23}$$

If the particles are electrons then each state can accommodate two—one with spin up and one with spin down—and the density of states $g_e(E)$ would be twice of what is given by Eq. (10.23).

$$g_e(E) = \frac{8\pi m V (2mE)^{\frac{1}{2}}}{h^3} \tag{10.24}$$

10.8 PLANCK'S BLACKBODY FORMULA

As an application of Bose-Einstein statistics, we now derive Planck formula for blackbody radiation. Radiation in thermal equilibrium inside a blackbody may be treated as a system of photons obeying BE statistics. Photons are bosons with spin I . In the case of molecular gas, the number of particles and the total energy remain constant. However, in this case photons are continually being created (emitted by the oscillators in the cavity wall) and destroyed (absorbed by the oscillators). Therefore, the number of photons in the system is not constant. Since the system is in thermal equilibrium, the total energy is constant. Consequently, the factor involving the undetermined multiplier α will not be there. Hence,

$$n_i = \frac{g_i}{\exp\left(\frac{E_i}{kT}\right) - 1} \tag{10.25}$$

Replacing E_i by hv , the number of particles dN in the frequency range v and $v + dv$ is given by

$$dN = \frac{g(v) dv}{\exp\left(\frac{hv}{kT}\right) - 1} \tag{10.26}$$

where $g(v) dv$ is the number of eigenstates in the frequency range v and $v + dv$. Proceeding as in Section 10.7.

$$g(v) dv = \frac{1}{h^3} \iiint_{p}^{p+dp} dx dy dz \iiint dp_x dp_y dp_z = \frac{4\pi V}{h^3} p^2 dp$$

where Eq. (10.19) is used and V is the volume of the black body. The momentum of the photon

$$p = \frac{h\nu}{c}, \quad dp = \frac{h}{c} d\nu$$

Hence,

$$g(\nu) d\nu = \frac{4\pi V}{h^3} \left(\frac{h\nu}{c} \right)^2 \left(\frac{h}{c} \right) d\nu = \frac{4\pi V}{c^3} \nu^2 d\nu \quad (10.27)$$

Since radiation of each frequency has two independent directions of polarizations, the above number has to be multiplied by two. This leads to

$$g(\nu) d\nu = \frac{8\pi V}{c^3} \nu^2 d\nu \quad (10.28)$$

Substituting this value of $g(\nu) d\nu$ in Eq. (10.26), we get

$$dN = \frac{8\pi V}{c^3} \frac{\nu^2 d\nu}{\exp\left(\frac{h\nu}{kT}\right) - 1} \quad (10.29)$$

If the number of photons per unit volume per unit frequency interval ($dN/V d\nu$) is multiplied by $h\nu$, what we get is the energy per unit volume per unit frequency interval which is the energy density u_ν . Therefore,

$$u_\nu = \frac{8\pi h\nu^3}{c^3} \frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1} \quad (10.30)$$

which is **Planck's blackbody formula**. This treatment shows the effectiveness of the statistical approach in deriving a very fundamental law.

10.9 ENERGY OF A FERMI-DIRAC SYSTEM

Electrons in a metal are loosely bound, and therefore, we can treat them as a gas of free particles, called **electron gas**, obeying F-D statistics. The F-D distribution law, Eq. (10.8) may be written in the form

$$dN = \frac{g(E) dE}{B_2 \exp\left(\frac{E}{kT}\right) + 1} \quad (10.31)$$

where dN is the number of electrons in the energy range E and $E + dE$, $g(E) dE$ is the number of eigenstates in the energy range E and $E + dE$, and $B_2 = e^\alpha$. The value of $g(E) dE$ for a free electron gas is given in Eq. (10.24). With this value of $g(E) dE$, Eq. (10.31) becomes

$$dN = \frac{8\pi mV}{h^3} \frac{(2mE)^{\frac{1}{2}} dE}{B_2 \exp\left(\frac{E}{kT}\right) + 1} \quad (10.32)$$

On integration, we get

$$N = \frac{8\pi m V (2m)^{\frac{1}{2}}}{h^3} \int_0^{\infty} \frac{E^{\frac{1}{2}} dE}{B_2 \exp\left(\frac{E}{kT}\right) + 1} \quad (10.33)$$

Total energy

$$E_T = \int E dN = \frac{8\pi m V (2m)^{\frac{1}{2}}}{h^3} \int_0^{\infty} \frac{E^{\frac{3}{2}} dE}{B_2 \exp\left(\frac{E}{kT}\right) + 1} \quad (10.34)$$

Evaluation of these integrals is difficult. It is convenient to write α as

$$\alpha = -\frac{E_F}{kT} \quad \text{or} \quad B_2 = e^\alpha = e^{-\frac{E_F}{kT}} \quad (10.35)$$

where E_F is the **Fermi energy** whose significance would be explained a little latter. Consequently, the F.D distribution function $f_{FD}(E)$ becomes

$$f_{FD}(E) = \frac{1}{e^{\frac{(E-E_F)}{kT}} + 1} \quad (10.36)$$

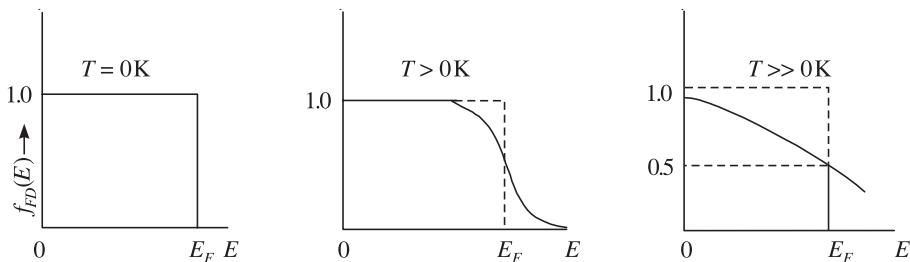
At $T = 0\text{K}$

$$\text{for } E > E_F : f_{FD}(E) = \frac{1}{e^{\infty} + 1} = 0 \quad (10.37)$$

At $T = 0\text{K}$, for $E < E_F$:

$$f_{FD}(E) = \frac{1}{e^{-\infty} + 1} = 1 \quad (10.38)$$

In other words, at $T = 0\text{K}$ all energy states from the ground state up to the Fermi energy are occupied and all states above the Fermi energy are empty. The situation is illustrated in Fig. 10.1 for (a) $T = 0\text{K}$ (b) $T > 0\text{K}$ (c) $T \gg 0\text{K}$



- (a) At $T = 0\text{K}$, all levels above E_F are unoccupied
- (b) For $T > 0\text{K}$, some particles near the Fermi energy can move to levels above E_F
- (c) For $T \gg 0\text{K}$ particles even in the ground state may move to higher levels.

Fig. 10.1 Fermi—Dirac distribution function $f_{FD}(E)$ for three different temperatures.

With the condition at $T = 0\text{ K}$, Eqs. (10.33) and (10.34) reduce to

$$N = \frac{8\pi m V (2m)^{\frac{1}{2}}}{h^3} \int_0^{E_F} E^{\frac{1}{2}} dE \quad (10.39)$$

$$E(0) = \frac{8\pi m V (2m)^{\frac{1}{2}}}{h^3} \int_0^{E_F} E^{\frac{3}{2}} dE \quad (10.40)$$

where $E(0)$ is the total energy at 0 K . Evaluating the integral in Eq. (10.39), we get

$$N = \frac{8\pi V}{3h^3} (2m)^{\frac{3}{2}} E_F^{\frac{3}{2}} \quad (10.41)$$

$$E_F = \frac{h^2}{2m} \left(\frac{3N}{8\pi V} \right)^{\frac{2}{3}} = \frac{h^2}{8m} \left(\frac{3N}{\pi V} \right)^{\frac{2}{3}} \quad (10.42)$$

Evaluation of E_F is straightforward as the ratio $\frac{N}{V}$ is easily obtained from Hall effect measurements. The Fermi temperature T_F and Fermi speed v_F are defined in terms of E_F by

$$T_F = \frac{E_F}{k}, \quad E_F = \frac{1}{2} m v_F^2 \quad (10.43)$$

Evaluation of the integral in Eq. (10.40) gives

$$E(0) = \frac{8\pi V}{5h^3} (2m)^{\frac{3}{2}} E_F \cdot E_F^{\frac{3}{2}} \quad (10.44)$$

Replacement of E_F by Eq. (10.42) leads to

$$E(0) = \frac{3Nh^2}{10m} \left(\frac{3N}{8\pi V} \right)^{\frac{2}{3}} \quad (10.45)$$

Replacement of $E_F^{\frac{3}{2}}$ in Eq. (10.44) using Eq. (10.42) gives

$$E(0) = \frac{3}{5} N E_F \quad (10.46)$$

Mean energy of the electron at absolute zero, $E_m(0)$, is given by

$$E_m(0) = \frac{3}{5} E_F \quad (10.47)$$

It may be noted that both classical and B.E statistics lead to a value of zero for the energy of a monatomic gas at $T = 0\text{ K}$. However, a Fermi-Dirac gas possesses appreciable energy at 0 K . This energy is called **zero point energy**.

Methods have been developed to evaluate the integrals in Eqs. (10.33) and (10.34) at temperatures above absolute zero. One of the expressions obtained for E_T is

$$E_T = \frac{3}{5} N E_F \left[1 + \frac{5\pi^2}{12} \left(\frac{kT}{E_F} \right)^2 + \dots \right] \quad (10.48)$$

The relation $C_v = \left(\frac{\partial E}{\partial T} \right)_v$ leads to

$$C_v = \frac{k^2 \pi^2 N T}{2 E_F} \quad (10.49)$$

That is, at very low temperatures the electronic contribution to the heat capacity is proportional to the temperature. In fact experimental observations support this trend.

SOLVED EXAMPLES

Example 10.1 The first excited state E_2 of the hydrogen atom is 10.2 eV above the ground state E_1 . The degeneracy of the ground and first excited states are 2 and 8, respectively. What is the ratio of the number of atoms in the first excited state to the number in the ground state at $T = 6000$ K.

Solution: The number of atoms in a state with energy E is given by Eq. (10.2). The required ratio is:

$$\frac{n_2}{n_1} = \frac{g_2}{g_1} e^{-\frac{(E_2 - E_1)}{kT}}$$

It is given that $g_2 = 8$, $g_1 = 2$, $(E_2 - E_1) = 10.2$ eV

$$kT = (1.38 \times 10^{-23} \text{ J/K}) (6000 \text{ K}) = 8.28 \times 10^{-20} \text{ J}$$

$$= \frac{8.28 \times 10^{-20} \text{ J}}{1.6 \times 10^{-19} \text{ J/eV}} = 0.52 \text{ eV}$$

$$e^{-\frac{(E_2 - E_1)}{kT}} = e^{-\frac{10.2}{0.52}} = e^{-19.2}$$

$$\text{Hence, } \frac{n_2}{n_1} = 4 \times e^{-19.2} = \frac{4}{3.32 \times 10^8} = 1.2 \times 10^{-8}$$

Example 10.2 What is the ground state energy of 10 non-interacting bosons in an one dimensional box of length a , if the energies of the various levels are given by $E_n = \frac{\pi^2 \hbar^2 n^2}{(2ma^2)}$, $n = 1, 2, 3, \dots$?

Solution: According to B-E statistics, an energy level can accommodate any number of bosons. Hence, all the 10 non-interacting bosons will be in the $n = 1$ state. The energy of the $n = 1$ state is given by

$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2}$$

The energy of the system of 10 bosons $= \frac{10\pi^2 \hbar^2}{2ma^2} = \frac{5\pi^2 \hbar^2}{ma^2}$

Example 10.3 Nine fermions are in an one-dimensional box of length a . Find the ground state energy of the system, if the energies of the various levels are given by $E_n = \frac{\pi^2 \hbar^2 n^2}{(2ma^2)}$, $n = 1, 2, 3, \dots$

Solution: According to F.D. statistics, an energy level can accommodate only 2 fermions, one with spin up and the other with spin down. Hence, two fermions will go in the level $n = 1$, two in the level $n = 2$, two in $n = 3$, two in $n = 4$ and the ninth fermion will be in the $n = 5$ level. Then the total ground state energy

$$\begin{aligned} E &= 2E_1 + 2E_2 + 2E_3 + 2E_4 + E_5 \\ &= 2 \frac{\pi^2 \hbar^2}{2ma^2} + 2 \frac{4\pi^2 \hbar^2}{2ma^2} + 2 \frac{9\pi^2 \hbar^2}{2ma^2} + 2 \frac{16\pi^2 \hbar^2}{2ma^2} + \frac{25\pi^2 \hbar^2}{2ma^2} \\ &= \frac{\pi^2 \hbar^2}{2ma^2} (2 + 8 + 18 + 32 + 25) = \frac{85\pi^2 \hbar^2}{2ma^2} \end{aligned}$$

Example 10.4 Density of silver is 10.5 g/cm³. Its atomic weight is 108 g/mol. If each atom contributes one electron for conduction, calculate (i) the number density of conduction electrons (ii) Fermi energy.

Solution: (i) Number density of conduction electrons is given by

$$\begin{aligned} \frac{N}{V} &= \frac{\rho N_A}{m} = \frac{(10.5 \text{ g/cm}^3)(6.02 \times 10^{23} \text{ atoms/mol})}{108 \text{ g/mol}} \\ &= 5.85 \times 10^{22} / \text{cm}^3 = 5.85 \times 10^{28} / \text{m}^3 \end{aligned}$$

(ii) Fermi energy is given by

$$\begin{aligned} E_F &= \frac{\hbar^2}{8m} \left(\frac{3N}{\pi V} \right)^{\frac{2}{3}} \\ E_F &= \frac{(6.626 \times 10^{-34} \text{ Js})^2}{8 (9.1 \times 10^{-31} \text{ kg})} \left(\frac{3}{\pi} \times 5.85 \times 10^{28} / \text{m}^3 \right)^{\frac{2}{3}} \\ &= 8.816 \times 10^{-19} \text{ J} = \frac{8.816 \times 10^{-19} \text{ J}}{1.6 \times 10^{-19} \text{ J/eV}} = 5.51 \text{ eV} \end{aligned}$$

Example 10.5 Calculate the electronic contribution to the molar heat capacity of sodium at 293 K. Fermi energy of sodium is 3.24 eV.

Solution: From Eq. (10.49), for heat capacity, we have

$$C_v = \frac{\pi^2 N k^2 T}{2 E_F}$$

where N is the total number of valence electrons. When we consider the molar specific heat, N becomes Avogadro's number.

$$\begin{aligned} C_v &= \frac{\pi^2 (6.02 \times 10^{23} / \text{mol}) (1.38 \times 10^{-23} \text{ J/K})^2 293 \text{ K}}{2 \times 3.24 \text{ eV} (1.6 \times 10^{-19} \text{ J/eV})} \\ &= 0.32/\text{mole} \end{aligned}$$

Example 10.6 The number density of conduction electrons in aluminium is $18.1 \times 10^{28} \text{ m}^{-3}$

(i) What is its Fermi energy? (ii) What is the mean energy of the electron at 0K?

Solution: (i) Fermi energy E_F is given by

$$\begin{aligned} E_F &= \frac{h^2}{8m} \left(\frac{3N}{\pi V} \right)^{\frac{2}{3}}, \quad \frac{N}{V} = 18.1 \times 10^{28} \text{ m}^{-3} \\ E_F &= \frac{(6.626 \times 10^{-34} \text{ Js})^2}{8(9.1 \times 10^{-31} \text{ kg})} \left(\frac{3 \times 18.1 \times 10^{28} \text{ m}^{-3}}{\pi} \right)^{\frac{2}{3}} \\ &= 18.698 \times 10^{-19} \text{ J} = \frac{18.698 \times 10^{-19}}{1.6 \times 10^{-19} \text{ J/eV}} \\ &= 11.69 \text{ eV} \end{aligned}$$

(ii) Mean energy of the electron at 0 K

$$E_m(0) = \frac{3}{5} E_F = 7.01 \text{ eV}$$

REVIEW QUESTIONS

- 10.1** What are micro states and macrostates?
- 10.2** State and explain the Boltzmann distribution function.
- 10.3** What is a phase space? For a system having f degrees of freedom, what is the dimension of volume in phase space?
- 10.4** Define density of states. State the expression for the number of eigenstates in the translational energy range E and $E + dE$.

- 10.5** What are the basic postulates of Maxwell-Boltzman statistics? Write down the distribution law for a system obeying Maxwell-Boltzman statistics.
- 10.6** State the basic postulates of Maxwell-Boltzman, Fermi-Dirac, and Bose-Einstein statistics.
- 10.7** Explain the distribution laws for Fermi-Dirac and Bose-Einstein statistics.
- 10.8** Write down the distribution formula for Maxwell-Boltzman, Bose-Einstein, and Fermi-Dirac statistics. When will all the three become equal?
- 10.9** Explain the condition for a system to obey Maxwell-Boltzman statistics.
- 10.10** What are bosons? Give two examples.
- 10.11** With examples explain what are fermions.
- 10.12** Classify the following into bosons and fermions: photons, protons, electrons, helium atoms.
- 10.13** Explain Fermi energy and Fermi speed. How are the two related?
- 10.14** Even at very low temperatures, fermions do not condense into the lowest energy levels. Why?
- 10.15** What is Fermi energy? Will Fermi energy depend on the number of particles in the system?
- 10.16** What are the basic postulates of Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics? State and explain the distribution law in all the three statistics.
- 10.17** What is a photon gas? Derive Planck's radiation law using Bose-Einstein statistics.
- 10.18** Derive an expression for the density of states for free electrons in a solid.
- 10.19** Using Fermi-Dirac distribution function deduce an expression for the Fermi energy of a system. Hence, obtain an expression for the mean energy of an electron at absolute zero.

PROBLEMS

1. The first excited rotational energy state of the H_2 molecule is about 4×10^{-3} eV above the ground state. The degeneracy of the ground and excited states are 1 and 3, respectively. What is the ratio of the numbers of these molecules in these two states at a temperature 300K?
2. Two energy levels of a system are separated by an energy of 8×10^{-21} J. The degeneracy of the lower state is 5 and that of the upper state is 9. What is the ratio of the number in the lower state to the number in the upper state at $T = 300$ K?
3. What is the ground state energy of 8 non-interacting fermions in a one dimensional box of length L , if the energy of the various levels are given by $E_n = \pi^2 \hbar^2 n^2 / (2ml^2)$, $n = 1, 2, 3, \dots$?

226 *Engineering Physics*

4. Six non-interacting bosons are confined in a one dimensional box of length l . If the energies of the levels are given by $E_n = \pi^2 \hbar^2 n^2 / (2ml^2)$, $n = 1, 2, 3, \dots$ what is the ground state energy of the system?
5. If number density N/V of conduction electrons in gold is $5.9 \times 10^{28} / \text{m}^3$, (i) What is the Fermi energy? (ii) What is the mean energy of the electron at 0K?
6. Calculate the Fermi energy of copper. The number density of conduction electrons in copper is $8.47 \times 10^{28}/\text{m}^3$.
7. Density of sodium is 0.97 g/cm³. Its atomic weight is 23 g/mol. If each atom contributes one electron for conduction, calculate the number density of conduction electrons in sodium, and the Fermi energy.
8. On the basis of Fermi-Dirac theory, compute the electronic contribution to the molar heat capacity of copper at 300K. Fermi energy of copper is 7.0 eV.

C H A P T E R

11

Lasers

Laser, perhaps one of the most exciting discoveries of the twentieth century, is an acronym for light amplification by stimulated emission of radiation. The first successful operation of a laser was done by Maiman in 1960 using a ruby crystal. Laser light, like light from any other ordinary source, is emitted when atoms make a transition from quantum state of higher energy to a state of lower energy. However, it has unique properties not found in the light from ordinary sources. Let us now discuss its basic principle, working of some laser systems, and its important applications.

11.1 ABSORPTION AND EMISSION OF RADIATION

11.1.1 Absorption

In any process of absorption or emission, at least two energy states are involved. Consider an atomic system with two energy states E_1 (lower) and E_2 . At ordinary temperatures most of the atoms will be in the lower energy state. When we allow electromagnetic radiation of frequency

$v = \frac{(E_2 - E_1)}{h}$ to interact with the system, some of the atoms will absorb the radiation and make

transitions from state 1 to state 2. This absorption of radiation of appropriate frequency by a system is called **stimulated absorption** or simply **absorption**. Symbolically, absorption is represented as:



Figure 11.1 illustrates this absorption process; the black spheres represent the atoms.

Atom* = Atom in its excited state.

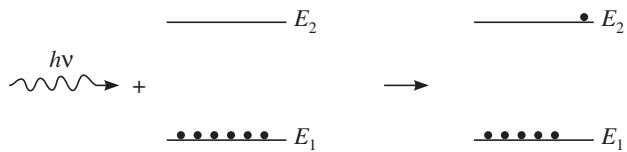


Fig. 11.1 Illustration of the absorption process.

11.1.2 Spontaneous Emission

In Fig. 11.2, we have two atoms in the excited state of energy E_2 and no external radiation is present. An atom in the excited state can emit a photon of energy $h\nu = E_2 - E_1$ on its own, and come to the lower energy state E_1 . The process of emission of photon without the influence of any external agency is called **spontaneous emission**, which is represented as follows:



Photons emitted spontaneously will not be in phase. Waves having a fixed phase relationship with each other are said to be coherent. Hence, light from an ordinary light source is incoherent.

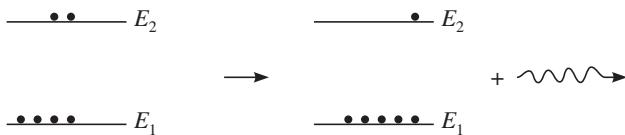
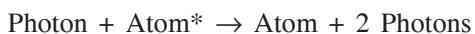


Fig. 11.2 Illustration of spontaneous emission.

11.1.3 Stimulated Emission

In Fig. 11.3 we have atoms in the excited state E_2 , but this time radiation of energy $h\nu = E_2 - E_1$ is also present. This photon can stimulate an atom in the excited state to emit a photon of energy $h\nu$. This process in which the emission is triggered by the external photon is called **stimulated emission** or **induced emission**. Symbolically,



The two photons will be moving in the same direction with the same phase. Hence, light output due to stimulated emission will be completely coherent. In what we considered here, the input is $1 h\nu$, whereas the output is $2 h\nu$. In other words, the stimulated emission process amplifies the intensity of the radiation.

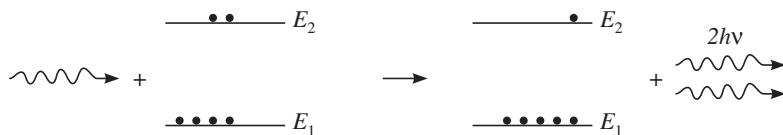


Fig. 11.3 Illustration of stimulated emission.

Atom* = Atom in its excited state.

11.1.4 Radiative and Non-radiative Transitions

Transitions between energy levels that occur with the absorption or emission of radiation are called **radiative transitions**. Transitions that occur without the absorption or emission of radiation are called **non-radiative transitions**. Non-radiative transition occurs mainly because of exchange of energy between the system and its surroundings. They are very common in laser materials.

11.2 PRINCIPLE OF LASER

In the previous section, we have seen that the photon emitted by a stimulated emission process and the photon that triggered the emission will be in phase, and will travel in the same direction. In a system having a large number of atoms, this process can occur many times, giving rise to a substantial amplification of the incident radiation. Lasers are devices that work on this principle of amplification by stimulated emission. If we have a collection of atoms in the excited state, the build up of an intense beam is illustrated in Fig. 11.4. At each stage, a photon interacts with an excited atom and produces two photons.

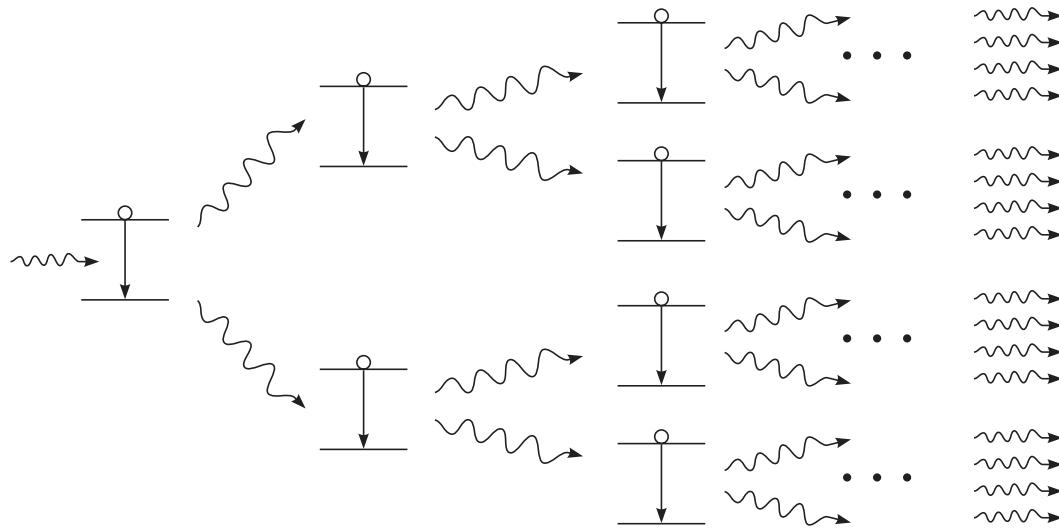
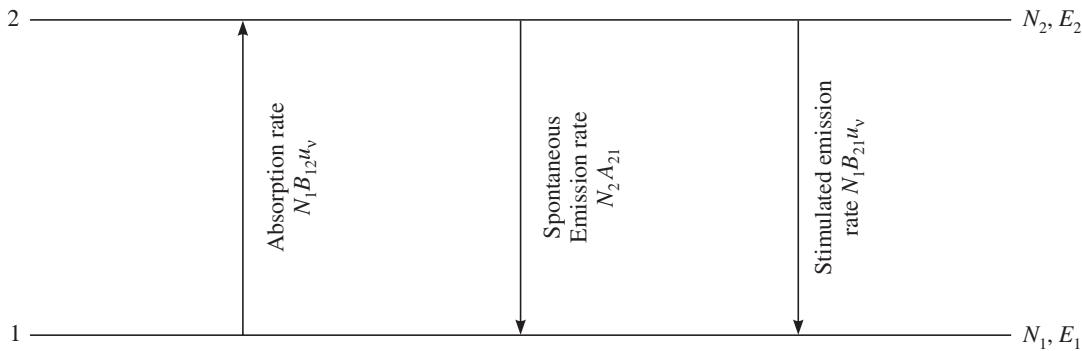


Fig. 11.4 Build up of an intense beam in a laser. Each photon interacts with an excited atom, giving rise to two photons.

11.3 EINSTEIN'S COEFFICIENTS

In the presence of radiation of appropriate frequency, in a system having a large number of atoms, all the three types of transitions—absorption, spontaneous emission and stimulated emission—occur simultaneously. For simplicity, consider a system having levels 1 and 2 (Fig. 11.5). Let their energies be E_1 and E_2 and $E_2 - E_1$ be $h\nu$.

**Fig. 11.5** Emission and absorption rates between states 1 and 2.

Let the total number of atoms be \$N_0\$. Under equilibrium conditions, let \$N_1\$ atoms be in state 1 and \$N_2\$ atoms be in state 2. By Boltzman distribution, we get

$$N_1 = N_0 e^{-\frac{E_1}{kT}} \quad \text{and} \quad N_2 = N_0 e^{-\frac{E_2}{kT}} \quad (11.1)$$

$$\frac{N_1}{N_2} = e^{\frac{(E_2 - E_1)}{kT}} = e^{\frac{h\nu}{kT}} \quad (11.2)$$

Let this system of atoms be in equilibrium with electromagnetic radiation of energy \$h\nu\$ and energy density \$u_v\$. The rate of absorption is proportional to the density of radiation \$u_v\$ and the number of atoms \$N_1\$ present in the lower state. That is, the

$$\text{Rate of absorption} = B_{12} N_1 u_v \quad (11.3)$$

where \$B_{12}\$ is the constant of proportionality, called the **Einstein's coefficient for absorption**. It can be shown that \$B_{12}\$ is simply the probability per unit time for absorption.

Next, we consider transitions from state 2 to state 1. Atoms in state 2 can come to state 1 both by spontaneous and stimulated emissions. The spontaneous emission rate depends only on the number of atoms \$N_2\$ in state 2:

$$\text{Rate of spontaneous emission} = A_{21} N_2 \quad (11.4)$$

The constant of proportionality \$A_{21}\$, called **Einstein's coefficient for spontaneous emission**, is the probability per unit time for spontaneous emission. The stimulated emission rate depends on the energy density \$u_v\$ of the radiation and the number of atoms \$N_2\$ in state 2. Hence,

$$\text{Rate of stimulated emission} = B_{21} N_2 u_v \quad (11.5)$$

where \$B_{21}\$ is **Einstein's coefficient for stimulated emission**.

When the system is in equilibrium, the rate of absorption from state 1 to 2 must be equal to the rate of emissions from state 2 to 1. That is,

$$B_{12} N_1 u_v = B_{21} N_2 u_v + A_{21} N_2 \quad (11.6)$$

$$u_v B_{21} N_2 \left(\frac{B_{12} N_1}{B_{21} N_2} - 1 \right) = A_{21} N_2$$

$$u_v = \frac{A_{21}}{B_{21} \left(\frac{B_{12} N_1}{B_{21} N_2} - 1 \right)} \quad (11.7)$$

Substituting the value of $\frac{N_1}{N_2}$ from Eq. (11.2), we have

$$u_v = \frac{A_{21}}{B_{21} \left(\frac{B_{12} e^{\frac{h\nu}{kT}}}{B_{21}} - 1 \right)} \quad (11.8)$$

From Planck's radiation law

$$u_v = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{kT}} - 1} \quad (11.9)$$

Comparing Eqs. (11.8) and (11.9), Einstein guessed that

$$B_{12} = B_{21} = B \quad (11.10)$$

$$\frac{A}{B} = \frac{8\pi h\nu^3}{c^3} \quad (11.11)$$

Equations (11.10) and (11.11) are called **Einstein's relations** and the coefficients are called **Einstein's A and B coefficients**. Equation (11.10) gives the important result that the probability for absorption from state 1 to 2 is equal to the probability for stimulated emission from state 2 to state 1. Since $B_{12} = B_{21}$ from Eq. (11.8), we get

$$\frac{A}{Bu_v} = e^{\frac{h\nu}{kT}} - 1 \quad (11.12)$$

From Eqs. (11.4), (11.5) and (11.12), we have

$$\frac{\text{Spontaneous emission rate}}{\text{Stimulated emission rate}} = \frac{A}{Bu_v} = e^{\frac{h\nu}{kT}} - 1 \quad (11.13)$$

If $h\nu \gg kT$, $e^{\frac{h\nu}{kT}} - 1$ will be very large and spontaneous emission far exceeds stimulated emission. At ordinary temperatures, this happens in the visible region. Stimulated emission becomes important when $h\nu \approx kT$, and may dominate when $h\nu \ll kT$, which happens in the microwave region. Hence, Townes tried the fabrication of **maser** (microwave amplification of stimulated emission of radiation) first.

11.4 POPULATION INVERSION

Again, consider an optical medium having states 1 and 2 through which a radiation of density u_v is passing. Let the number of atoms in the lower and upper energy states be N_1 and N_2 , respectively. Under equilibrium condition, according to Boltzman distribution, $N_1 > N_2$ as shown in Fig. 11.6(a).

$$\text{The absorption rate} = N_1 Bu_v$$

$$\text{The stimulated emission rate} = N_2 Bu_v$$

where B is Einstein's B coefficient. The stimulated emission rate will exceed the absorption rate only if $N_2 > N_1$. Such a situation [Fig.11.6(b)] is called **population inversion**. The situation for which $N_2 > N_1$ can still be represented by the Boltzman condition, Eq. (11.2), if we assume that T is negative. Therefore, the condition $N_2 > N_1$ is often referred to as one of **negative temperature**. Hence, amplification is possible only if we generate a population inversion by artificial means. When a two level system is irradiated by a radiation of frequency

$$v = \frac{(E_2 - E_1)}{h}, \text{ the number of atoms in level 2 increases until } N_2 = N_1. \text{ Further increase is not}$$

possible since at that stage the rate of absorption equals the rate of stimulated emission. Population inversion is therefore not possible in two level systems. Hence, for light amplification we require a system having at least three levels. Population inversion in three and four level systems are discussed in Section 11.7.

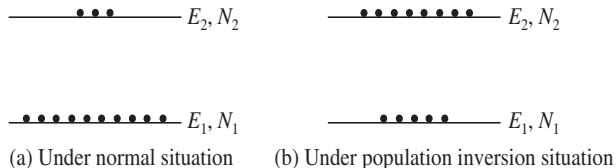


Fig.11.6 Two level systems.

11.5 BASIC COMPONENTS OF A LASER

A laser requires three basic components for its operation:

- (i) An **active medium** to amplify the light.
- (ii) An energy source, called **pump source**, that excites the active medium to excited states, which results in the creation of population inversion in the medium.
- (iii) A resonant cavity along with the active medium inside to provide the necessary feedback for the production of a very intense beam.

In most of the lasers the active medium is enclosed in the resonant cavity formed by two mirrors facing each other. One of the mirrors is less than 100% reflective. The arrangement of the basic components is illustrated in Fig. 11.7. The pump source excites the active medium and creates the necessary population inversion in the medium. A few spontaneously emitted

quanta which are available in the medium trigger the in-phase release of more quanta. The process continues and an avalanche action takes place. When the radiation is directed back and forth through the medium by these mirrors, the avalanche action continues, resulting in a very intense beam. As one of the mirrors is less than 100% reflective, the laser beam comes out through that mirror.

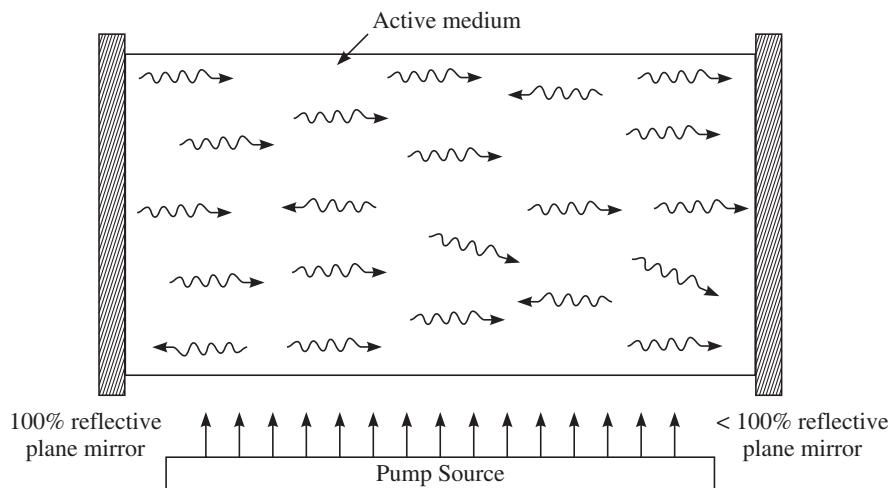


Fig. 11.7 The basic components of a laser.

11.5.1 Active Medium

The active medium is the one that amplifies the light. It may be a solid, liquid or gas. A synthetic ruby rod was used as the active medium in the first laser. A number of commercial lasers are available with Nd^{3+} ion as the active element in different host materials (Nd^{3+} in YAG, Nd^{3+} in glass, etc.). Another common solid active medium is Ti^{3+} ion in Al_2O_3 . Early liquid lasers contained chelates and metallo organic compounds such as europium benzoyl acetone. Most of the present day liquid lasers are the dye lasers. Dyes are solids that are dissolved in solvents like water, ethyl alcohol, and methyl alcohol. Some of the common dyes are rhodamine 6 G, rhodamine B, puronin B, etc. The active element in gas phase lasers can be a noble gas atom ($\text{He} - \text{Ne}$), a positive ion (Ar^+), a neutral molecule (N_2, CO_2), a metal atom (Cu vapour, $\text{He} - \text{Cd}$), and so on. Some of the gas lasers can give light even in the ultraviolet region.

11.5.2 Pumping Methods

The process of raising the atoms or molecules of the active medium to higher energy states is called **pumping**. The commonly used pumping methods are:

- | | |
|-------------------------------------|---------------------------|
| (i) Optical pumping | (ii) Electrical discharge |
| (iii) Inelastic atom-atom collision | (iv) Chemical reaction |
| (v) Direct conversion | |

Of these, the first two are used widely. In a chemical laser, the energy required to pump a laser comes from a chemical reaction. In direct conversion, electrical energy is converted into radiation, which occurs in LED's and in semiconductor lasers.

Optical pumping

Optical pumping was the method used in the first laser that was fabricated by Maiman. In this procedure a high intensity light source such as a flash lamp is used to excite the active medium. The output of the flash lamp is usually broad. Since the energy levels of the active medium in solid state lasers are broad, they can easily absorb the flash lamp output. Pumping by flash lamp is not suitable in gas lasers as their energy levels are fairly sharp. To increase the efficiency, the active medium is placed inside a helical flash lamp.

Electric discharge

The extremely high electric field in an electrical discharge accelerates the electrons emitted by the cathode towards the anode. Collisions between these high-energy electrons and the active medium produce atoms, molecules and ions of the medium in the excited states, producing the needed population inversion. This procedure is mainly used in gas lasers.

Inelastic atom-atom collision

This procedure is suitable when we have two types of atoms in the active medium. An electric discharge raises one type of atoms to their excited states. These excited atoms collide inelastically with the second type of atoms, energy exchange takes place, and the required population inversion is created in the later atoms. An example of this type is the He - Ne laser.

11.5.3 Resonant Cavity

In a laser, to generate high intensity outputs the light is directed back and forth through the medium. For this the active medium is placed between two mirrors. Such an arrangement is called a **resonator**. Resonator cavity of different forms are in use. Some of them are discussed below.

Plane parallel resonator

This arrangement has two parallel plane mirrors placed on either side of the medium, as shown in Fig. 11.7. A very precise alignment of this configuration is difficult. The radiation field generates a standing wave pattern that has nodes at both ends. The cavity will be resonant for the wavelengths given by

$$L = p \frac{\lambda}{2}; \quad p = 1, 2, 3, \dots \quad (11.14)$$

where L is the length of the cavity and λ is the wavelength. Since we have the active medium between the mirrors, L has to be replaced by the optical path length Ln where n is the refractive index of the medium. Hence, in terms of frequency, we have

$$v = p \frac{c}{2Ln}; \quad p = 1, 2, 3, \dots \quad (11.15)$$

These different frequencies, called the **axial modes** lie within the width of a single emission line. The output of a laser, therefore, need not be exactly monochromatic.

Confocal cavity

This arrangement has two identical concave mirrors separated by a distance equal to their radius of curvature R .

Spherical resonator

It consists of two concave mirrors of the same radius of curvature R separated by a distance $2R$. Sometimes, this is called the **concentric resonator**.

Long radius cavity

This is the most commonly used arrangement. It uses two concave mirrors of equal and fairly long radius of curvature R separated by a distance less than R .

Hemispherical cavity

It consists of a plane mirror at one side and a concave spherical mirror at the other. The plane mirror is placed at the centre of curvature of the spherical mirror. This is considerably easier to align than the others. However, the power output may be less.

11.6 METASTABLE STATES

An atom or molecule in an excited state remains there for a certain time called the **lifetime of that state**, before making a transition to a lower state. The lifetime of a state is characteristic of the energy level and varies over a wide range. Most of the states have a short lifetime, of the order of 10^{-8} s. However, some energy states have very long lifetime, of the order of 10^{-3} s or higher. Energy states with such long lifetimes are called **metastable states**. The occurrence of metastable states may be explained as follows. In certain cases, regular selection rules do not permit transitions from a particular state to a lower state. In such cases the system has to remain in that state for longer time until weak perturbations and stimulating radiations induce transitions to lower state. Existence of metastable states are of fundamental importance in lasers.

11.7 THREE AND FOUR LEVEL SYSTEMS

11.7.1 Three Level System

We have already seen that for light amplification we require a system having at least three levels (Section 11.4). Let us consider the three level system shown in Fig. 11.8(a). The atoms originally in the ground state (E_1) are pumped to a higher state (E_3) by an external source of

energy. Rapid spontaneous emissions bring the excited atoms down to level E_2 , which is a metastable state. Hence, there will be a build up of atoms in level E_2 , resulting in a population inversion. The spontaneously emitted photons available in the system may then stimulate the emission of photons of same wavelength moving in the same direction with the same phase.

There is an inherent drawback with the three level system discussed here. Since the lower laser level is the ground state, to achieve population inversion we have to transfer more than half of the atoms to level 2. For that, very high pump power is required. To reduce the pump power one has to avoid the use of the ground state as the lower laser level. There is yet another difficulty with a three level system. As lasing starts, the population in the ground state increases. As a result, it is possible that the photons emitted by stimulated emission are reabsorbed, thereby removing photons that might contribute to the lasing action. We can get over these hurdles by using a four level system.

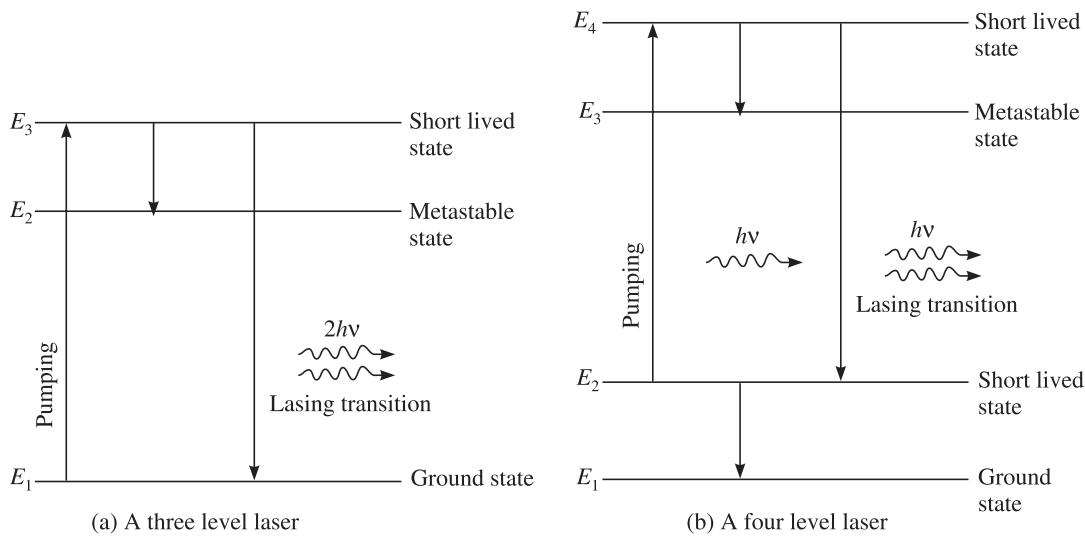


Fig. 11.8 Transitions in three level and four level systems.

11.7.2 Four Level System

A four level laser system is illustrated in Fig. 11.8(b). Atoms are pumped from the ground state to an excited state E_4 where they decay rapidly to the metastable state E_3 . The lasing transition proceeds from the metastable state to another excited state E_2 . Atoms in E_2 decay rapidly to the ground state. Since the lower laser level decays rapidly, population inversion is maintained efficiently. Moreover, as the lower laser level is not the ground state, atoms there cannot absorb the radiations emitted by stimulated emission.

11.8 RUBY LASER

Ruby laser is a three level solid state laser. It was the very first laser fabricated by Maiman in 1960. Ruby is a crystal of Al_2O_3 doped with Cr^{3+} ions. The impurity Cr^{3+} is the active material.

Typical Cr^{3+} concentrations are ~ 0.5 per cent by weight. The ruby is made into a cylindrical rod of 10 to 15 cm long and about 5 mm in diameter. One end of the rod is completely polished while the other end is partially reflecting. Therefore, the ends act as the cavity mirrors. The energy levels of the Cr^{3+} ions responsible for the laser action are shown in Fig. 11.9. The lower laser level of the system is the ground state G (level 1). The upper laser level is the metastable state M (level 2), which has a lifetime of about 10^{-3} s. Level 3 of the system consists of two bands F_1 and F_2 .

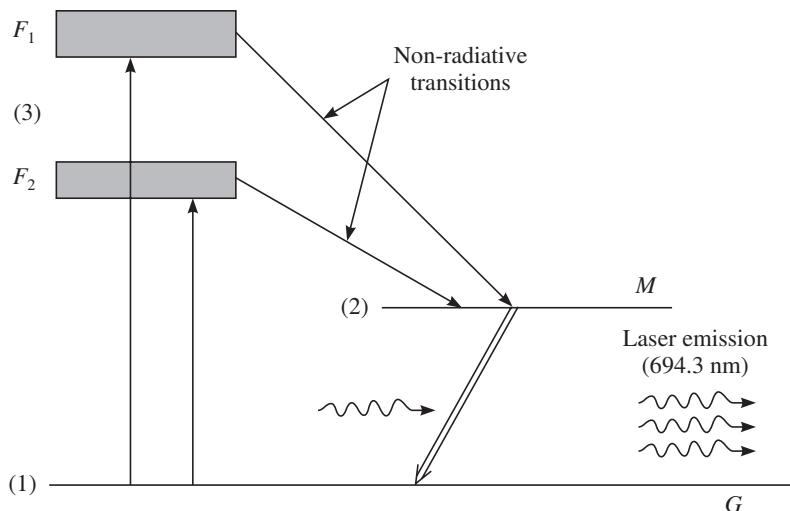
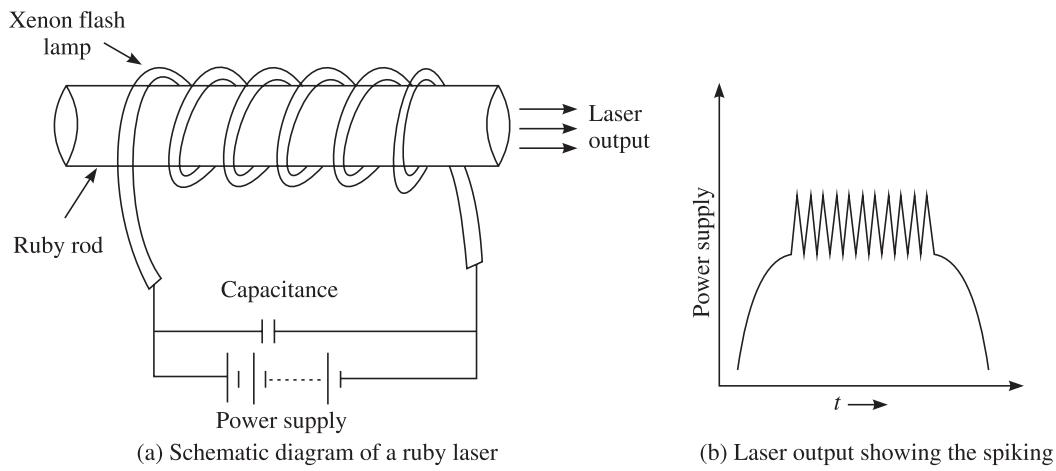


Fig. 11.9 Energy levels of chromium ions in the ruby laser and the involved transitions, where G is the ground state and M is the metastable state.

The states in the F_1 and F_2 bands have extremely small lifetimes, of the order of 10^{-9} s. The rod is placed inside a xenon flash lamp as shown in Fig. 11.10(a).

The light from xenon flash lamp pumps the Cr^{3+} to states in the energy bands F_1 and F_2 . The ions in these bands make a very fast non-radiative decay to the metastable state M . As this state has a longer lifetime, the number of ions in this state keeps on increasing, resulting in a population inversion between levels 1 and 2. When the required population inversion is reached, the spontaneously emitted photons existing in the system trigger the laser action in the medium. The output (694.3 nm) is pulsed since flash lamp operation is pulsed. The pumping flash lasts for a time duration of about 1 ms. Once the flash lamp is fired, within a very short time the population inversion exceeds the threshold value. Hence, stimulated emission starts and depopulates the upper laser level much faster than the pump rate. The laser action therefore stops momentarily and before the laser output falls to zero, the population inversion again exceeds the threshold value, thereby initiating laser action. This process is repeated many times before the pumping flash ends. Hence, the output consists of a large number of spikes of about 1 μs duration as shown in Fig. 11.10(b). Very high peak power pulses can be obtained by Q-switching.

**Fig. 11.10** Ruby laser schematic and output.

Since the lower laser level is the ground level, about one half of the ions have to be excited to obtain population inversion. Hence, for operation, a high pump power is required. As a lot of energy gets dissipated in the ruby rod, it has to be cooled for continuous operation.

11.8.1 Applications

Ruby laser is used for different types of applications. Some of them are as follows:

- (i) Measurement of distances using pulse echo techniques.
- (ii) For drilling holes of small l/d ratio (where l is length and d is diameter) in hard metals.
- (iii) For trimming resistors and integrated circuit masks.
- (iv) For the welding of detached retina, to destroy tissue in a localized area, for the treatment of skin disorders, etc.
- (v) In pulsed holography.
- (vi) In scientific research such as plasma production, study on fluorescence, etc.

11.9 HE-NE LASER

Helium-neon laser was the first gas laser fabricated by Javan, in 1961. It is an atomic laser. It gives a low power output, of a few milliamperes. He-Ne lasers giving laser light at 3391.3 nm, 1152.3 nm, and 632.8 nm are available commercially. A typical laser consists of a tube 80 to 90 cm in length and 1 cm in diameter, with electrodes on the sides and silica windows set at Brewster's angle. It is filled with 1 mm mercury of He and 0.1 mm mercury of Ne. The electrodes are connected to a high voltage source. The Ne atoms provide the energy levels for the laser transitions. He atoms provide an efficient mechanism for the excitation of neon atoms. The resonator mirrors are placed outside the tube, as shown in Fig. 11.11.

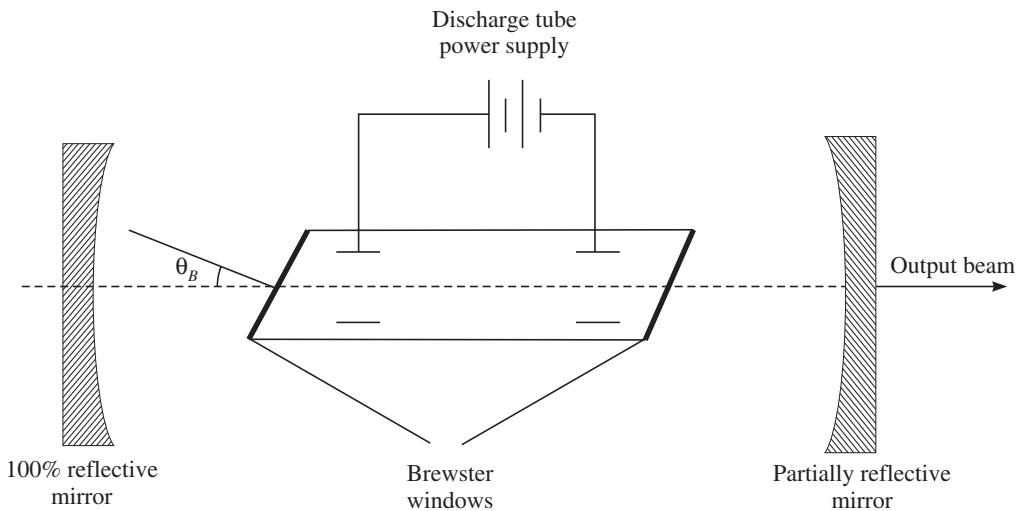
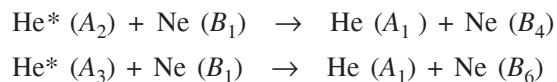


Fig. 11.11 Typical He—Ne laser with external mirrors.

To understand the excitation and laser emission processes, the energy levels of He and Ne are needed which is illustrated in Fig. 11.12.

When the discharge process starts, high energy electrons are produced in the tube. These collide with the atoms of He and Ne and excite them to higher energy states. Collisions with He atoms are more probable as their concentration is higher. Hence, a substantial number of He atoms are excited to a variety of excited states. In the normal cascade of these to the ground state, many collect in the metastable states A_2 and A_3 . These states nearly coincide in energy with the levels B_4 and B_6 of Ne atom. The lifetime of the states A_2 and A_3 are higher than the time between collisions of a He atom with a Ne atom. Helium atoms in levels A_2 and A_3 then collide with neon atoms in the ground level B_1 and exchange the energy. This results in the excitation of the Ne atoms to the levels B_4 and B_6 and the de-excitation of the He atoms to the ground level A_1 . Symbolically,



where the letter in parenthesis refers to the corresponding energy level. Direct excitation of the Ne atoms to B_4 and B_6 is also possible which will be much less, as neon pressure is less. The neon states are relatively metastable, and therefore, quickly reach population inversion relative to the lower states. Details regarding the three major laser transitions are given below:

The 632.8 nm transition

This transition corresponds to emission from $B_6 \rightarrow B_3$. The lifetime of B_3 is of the order of 10^{-8} s and that of B_6 is 10^{-7} s. Hence, population build up is possible. The lower laser level B_3 decays spontaneously into the level B_2 . This He-Ne laser light is the most widely used one.

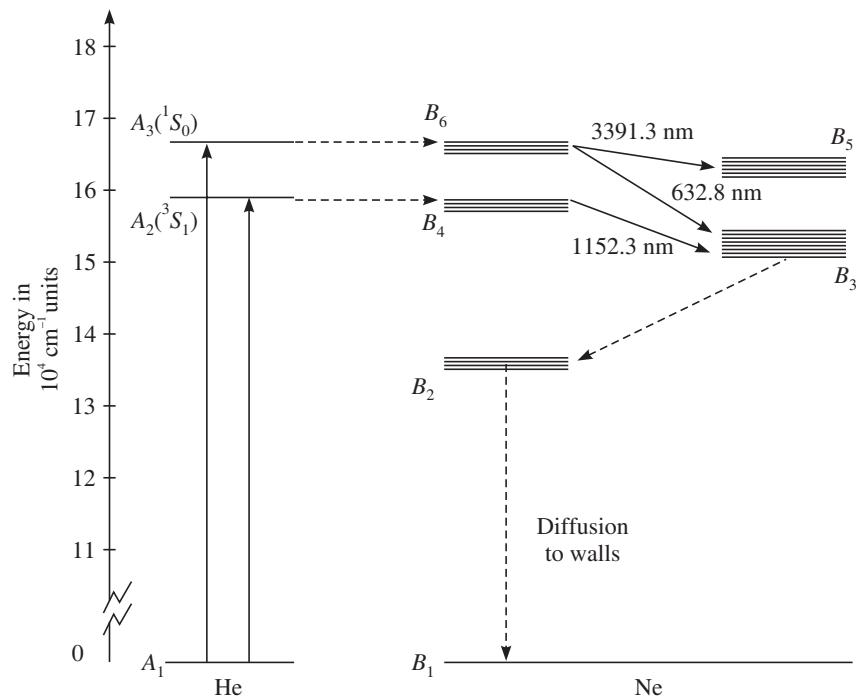


Fig. 11.12 He-Ne energy levels and the laser transitions.

The 1152.3 nm transition

This was the output wavelength of the first He-Ne laser. This transition corresponds to emission from $B_4 \rightarrow B_3$. The lower laser level of this transition is the same as that of the 632.8 nm one. The gain of 1152.3 nm and 632.8 nm transitions increases with decreasing diameter of the tube.

The 3391.3 nm transition

This transition corresponds to emission from $B_6 \rightarrow B_5$. The upper laser level is the same as the upper one of 632.8 nm transition. Although the upper laser level of the two are the same, the 3391.3 nm usually occurs first because of the shorter lifetime of level B_5 .

Some of the important applications of He-Ne lasers are listed as follows:

- (i) Extensively used in laboratories where highly coherent and monochromatic sources are needed.
- (ii) Widely used in devices such as supermarket scanners for barcode reading, printers, image and pattern recognition set ups.
- (iii) As range finders.
- (iv) In fourier transform spectrometers.
- (v) For recording and reconstruction of holograms.
- (vi) For non-destructive testing, including surface flaw and roughness measurements.
- (vii) To align other lasers.

11.10 CARBON DIOXIDE LASER

Lasers described so far are based on transitions between electronic states. The CO₂ laser is a molecular laser in which the involved energy levels are states representing the vibrational motion of atoms in the molecule. The CO₂ molecule is linear and will have 4 normal vibrations. They are depicted in Fig. 11.13. The first and second vibrations are completely symmetric. Hence, they are called the symmetric stretching and symmetric bending modes. The third one is the asymmetric stretching mode. The vibrations are labelled as v₁, v₂, and v₃ (Fig. 11.13).

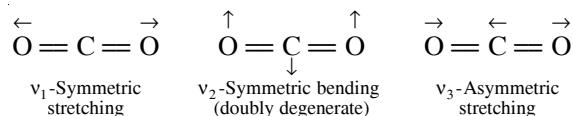


Fig. 11.13 The normal vibrations of CO₂ molecule.

The solid arrows indicate the direction of motion of atoms at a particular instant. The energy difference between the vibrational levels fall in the infrared region. The vibrational energy of a harmonic oscillator is given by

$$E_v = \left(v + \frac{1}{2}\right)hv, \quad v = 0, 1, 2, \dots \quad (11.16)$$

The energy of the CO₂ molecule is then

$$E(v_1, v_2, v_3) = \left(v_1 + \frac{1}{2}\right)hv_1 + \left(v_2 + \frac{1}{2}\right)hv_2 + \left(v_3 + \frac{1}{2}\right)hv_3 \quad (11.17)$$

where v₁, v₂, v₃ = 0, 1, 2, ... and v₁, v₂, v₃ are the frequencies of the 3 normal modes. Fig. 11.14 shows some of the low lying vibrational levels of CO₂. The level (000) means the ground state, (100) means the first excited state of symmetric stretching mode, (020) means the second excited state of the symmetric bending mode, and so on.

The population inversion is achieved in a discharge tube which contains CO₂ gas. Addition of N₂ or N₂ and He increases the efficiency of the laser. The vibrational states v = 0 and v = 1 of N₂ molecule are also included in Fig. 11.14. The CO₂ molecules excited by electron impact cascade down from their original level of excitation. A sizable fraction tends to collect in the long lived (001) level of CO₂. This creates a population inversion, which results in laser transitions at 10.6 μm and 9.6 μm (Fig. 11.14). When N₂ is present, a very large fraction of the excited N₂ molecules tend to collect in its v = 1 level, which coincides with the (001) level of CO₂. Collision of these N₂ molecules with the ground state CO₂ molecules excite the CO₂ molecules to the (001) state. At the temperatures and pressures involved in the operation of CO₂ laser, most of the N₂ molecules in the v = 1 state decay by this process. Thus, the addition of N₂ increases the population inversion by a large factor. In other words, the addition of N₂ increases the efficiency of the laser.

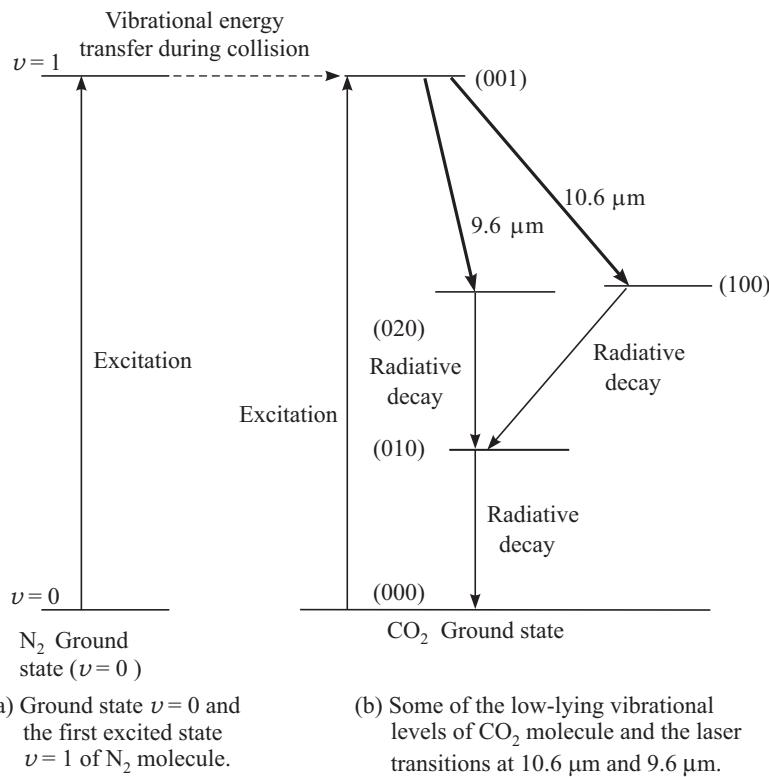


Fig. 11.14 N_2 and CO_2 vibrational levels and laser transitions.

Laboratory size CO_2 lasers can give an output of a few kilowatts. High power lasers are extensively used for welding, cutting, drilling, etc. The wavelength of the CO_2 laser output falls in a region where atmospheric attenuation is negligible. Hence, these lasers find applications in optical radar systems.

11.11 SEMICONDUCTOR LASER

When a p-n junction, having heavily doped p and n regions, is forward biased by connecting the p-material to the positive and the n-material to the negative of a direct current source, the electrons and holes combine across the depletion region to produce radiation. The photons thus produced may interact (i) either with the conduction band electrons stimulating radiative recombination (stimulated emission) thereby producing further photons or (ii) with valence band electrons and get absorbed (stimulated absorption). If the voltage across the p-n junction is large, the injected carrier concentration would be large. In such a case, the stimulated emission can exceed the absorption, producing amplification in the active region, which results in a good light output.

Figure 11.15 shows a schematic diagram of a p-n junction laser, the shaded region being the depletion region, usually of the order of $0.1 \mu\text{m}$. A resonant cavity is formed by just cleaving

the junction ends. The two opposite end faces are polished flat and parallel. The light that is trapped is reflected back and forth between these two faces. The other two sides are rough, and therefore, suppress oscillations in unwanted directions. Since the refractive index of a semiconductor crystal is about 3.6, the reflectivity of the crystal air interface is fairly high (about 32%). If needed, the surfaces may be coated to reduce the threshold current.

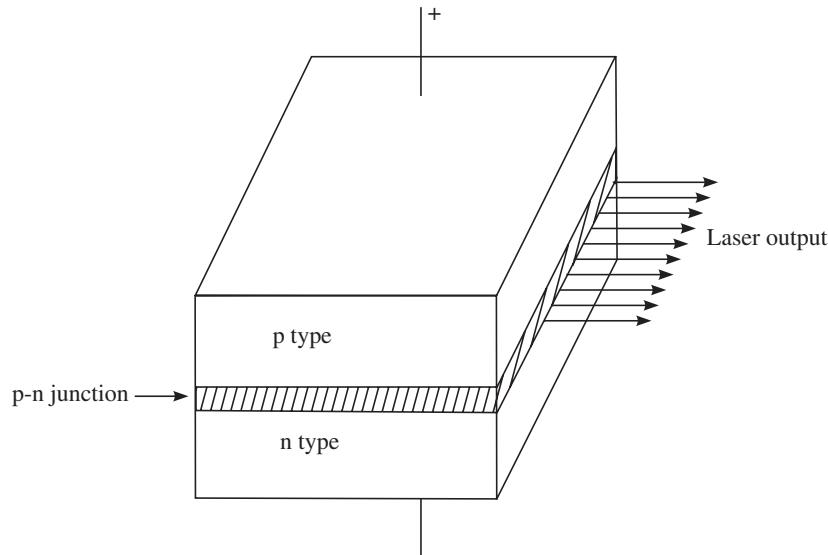


Fig. 11.15 Representation of a homojunction semiconductor laser.

The active region is a little wider than the depletion region. The transverse dimension of the beam is much greater than the width of the active region, the output beam ends up with a large beam divergence. Moreover, the output is not as monochromatic as that from a gas laser. A commonly used semiconductor is GaAs. By selecting suitable semiconducting materials and dopants, diode lasers emitting radiations from UV to IR can be fabricated. These lasers work continuously only at low temperatures. At room temperatures, the operation is usually pulsed. However, with the use of heterojunction diodes, continuous room temperature operation is possible. With multilayers of laser structure, output powers of about 10 mW can be produced with operating currents of less than 50 mA. Lasers with lifetimes of 40,000 h are now available.

The high efficiency, the simple working method, the miniature size, and the possibility of direct modulation make semiconductor lasers very attractive for a number of applications, optical communication in particular. The emission wavelength of GaAs, GaInAsP,...etc. matches with the transmission windows where optical fibers show minimum attenuation. Semiconductor lasers are also used in laser printers, supermarket scanners, optical storage drives, CD players, and CD rom drives.

11.12 CHARACTERISTICS OF LASER LIGHT

Like any other radiation, laser light is electromagnetic in nature. It differs from the conventional one in a number of ways as follows:

- (i) Highly monochromatic
- (ii) Highly coherent
- (iii) Highly directional
- (iv) Intense

11.12.1 Monochromaticity

In general, the light sources are not strictly monochromatic, but has a frequency width Δv . For conventional light sources Δv is of the order of 10^{10} Hz, whereas for a good laser it is of the order of about 500 Hz. The monochromaticity of a source is defined by the quantity $(\Delta v/v)$. For a source having wavelength $\lambda = 5000 \text{ \AA}$, $v = 6 \times 10^{14}$ Hz.

For a conventional source,

$$\frac{\Delta v}{v} = \frac{10^{10} \text{ Hz}}{6 \times 10^{14} \text{ Hz}} = 1.7 \times 10^{-5}$$

For a laser source,

$$\frac{\Delta v}{v} = \frac{500 \text{ Hz}}{6 \times 10^{14} \text{ Hz}} = 8 \times 10^{-13}$$

which is extremely small compared to that of the conventional source.

11.12.2 Coherence

Coherence means that two or more waves in a radiation field bear the same phase relationship to each other at all times. Light emitted by ordinary light sources is completely incoherent. However, light waves from lasers are in phase with each other, and travel in the same direction as they are due to stimulated emission. The coherence length (section 3.4) of a well-controlled laser is of the order of a few kilometers, which implies that the laser is highly coherent.

11.12.3 Directionality

The output beam of a laser has a well-defined wavefront, and therefore, it is highly directional except for the divergence caused by diffraction effects. If a beam emerges from a lens of diameter D , the angular spread due to diffraction is:

$$\theta = \frac{1.22\lambda}{D} \quad (11.18)$$

The high directionality allows us to focus it into a point by passing the beam through a suitable convex lens. If f is the focal length of the lens and a is the radius of the beam, then the area of the spot A at the focal plane is given by

$$A = \frac{\pi\lambda^2 f^2}{a^2} \quad (11.19)$$

where λ is the wavelength of light. Typically, if $a = 2$ mm, $\lambda = 5000$ Å and $f = 5$ cm.

$$A = \frac{\pi\lambda^2 f^2}{a^2} = \frac{\pi(5 \times 10^{-7} \text{ m})^2 (0.05 \text{ m})^2}{(2 \times 10^{-3} \text{ m})^2} = 4.9 \times 10^{-10} \text{ m}^2$$

which is extremely small.

11.12.4 Intensity

Laser output is highly intense. Since the laser power is concentrated in a beam of a few millimeters, even a small laser can deliver very high intensity at the focal plane of the lens. If the power of the laser is P Watts, then the intensity I at the focus is:

$$I = \frac{P}{A} = \frac{Pa^2}{\pi\lambda^2 f^2} \quad (11.20)$$

If P is 1 Watt, then the intensity at the focus for the above values is given by

$$I = \frac{1 \text{ W}}{4.9 \times 10^{-10} \text{ m}^2} = 2.04 \times 10^9 \text{ W/m}^2$$

11.13 APPLICATIONS OF LASERS

Lasers offer a variety of applications and in this section we shall discuss some of them.

11.13.1 Lasers in Industry

The large intensity that is possible in the focussed output of a laser beam and its directionality makes laser an extremely useful tool for a variety of industrial applications. Table 11.1 lists some of the fabrication processes, the nature of laser used, and the involved materials. Of the number of lasers available, CO₂ laser is the most commonly used since pulsed operation and continuous wave operation are possible. Capital cost is also reasonable.

Table 11.1 Fabrication processes and the nature of laser used

Process	Materials involved	Laser output used
Profile cutting	Metals, plastics, ceramics, glass, composites, textiles, wood	CW
Welding	Metals, plastics, composites	CW, Pulsed
Drilling	Metals, ceramics, plastics	Pulsed
Machining	Metals, plastics, ceramics	CW, Pulsed
Melting	Metals, non-metals, refractive materials	CW
Scribing	Ceramics	CW, Pulsed

Laser cutting

The cutting process is essentially removal of the material from the cut. The effectiveness of a laser for cutting can be increased by the use of a gas jet coaxial with the laser beam. The depth of the cut increases as the pressure of the gas jet is increased. The gas jet also helps in removing metal vapour and particles. Lasers are used extensively for cutting sheet metals and alloys of about 5 mm or less thickness in aircraft and automobile industry.

Welding

In welding, filler material is added to join two components. The source we use should not evaporate the material. The laser used for welding must have a high average power than high peak power. Neodymium lasers are preferred to ruby lasers, as the longer pulse width helps overall heat transfer without excessive vaporisation. Continuous wave CO₂ lasers have been used for deep penetration welding in which heat is transferred along the depth of the weld. For welding in thin sheets and in foils where the weld region is wider than its depth, pulsed CO₂ lasers are preferred.

Drilling

Vaporisation of material at the focus of the laser beam is the principle behind its use in drilling. A laser pulse having an energy of about 0.05 J and pulse width of 10⁻³ s can drill a hole of about 0.1 nm radius in a steel plate of thickness 1 mm. Pulsed ruby and neodymium lasers are commonly used for drilling holes of small l/d ratio in hard metals, which is difficult by other methods. Here, l is the thickness of the work and d is the diameter of the hole. Elastomeric materials are difficult to perforate by normal methods as they readily deform. Nowadays, lasers are extensively used wherever the material shows a tendency to deform during drilling. Lasers are used even to perforate teats for baby bottles.

Manufacture of electronic components

Electronic industry is considered among the first to use lasers in production applications. Laser systems have been employed for cutting, welding, and controlled evaporation during manufacture of microcircuits and integrated circuits. Lasers are used to perforate and divide silicon slices having several hundred integrated circuits. It is also used for the isolation of faulty components in a large integrated circuit by disconnecting the conducting paths by evaporation. Yet another important application is in the production of masks for integrated circuits. Trimming of thick and thin film resistors using lasers is a very common feature. Contact free operation, accuracy of positioning, ease of automation, and high output obtainable are some of the advantages.

11.13.2 Lasers in Scientific Research

With the advent of different types of lasers, they are extensively used as coherent, monochromatic, and intense light sources in different branches of science. Every laser that is developed has been utilised in an experiment of one form or another. Its use in spectroscopic experiments is so enormous that a new branch of spectroscopy, called **laser spectroscopy**, has now emerged. The use of lasers as an excitation source for the study of Raman effect has

completely revolutionised the applications of Raman spectroscopy as an analytical tool. Raman microscopy using lasers is becoming a very powerful microscopic method. A few of the important applications are mentioned below.

Study of nonlinear phenomena

When electromagnetic radiation is allowed to interact with a medium, it is polarized. The polarization \mathbf{P} is given by

$$\mathbf{P} = \chi_1 \mathbf{E} + \chi_2 \mathbf{E}^2 + \chi_3 \mathbf{E}^3 + \dots +$$

where \mathbf{E} is the electric field associated with the radiation, χ_1 is a constant called dielectric susceptibility, and χ_2, χ_3, \dots , are the higher order susceptibilities. With normal light sources, one would be able to study only the contribution from the first term which is linear in \mathbf{E} . The contribution from the higher order terms are referred to as **nonlinear effects** because of the nonlinear dependence on \mathbf{E} . The nonlinear contributions are observable only if the electric field \mathbf{E} is extremely high. Such a high electric field can only be produced by giant pulse lasers as they are capable of producing power densities of the order of 10^{10} to 10^{11} W/cm². Thus, high power lasers are finding important applications in observing nonlinear effects such as harmonic generation, observation of multi-photon processes, hyper Raman scattering, stimulated Raman scattering, etc.

Harmonic generation

Fixed frequency lasers emit coherent radiation at a single frequency or at a limited number of discrete frequencies. Dye lasers operate at wavelengths ranging from 1.5 nm to 350 nm. Harmonic generation ($2\nu_1, 3\nu_1, \dots$) and generation of sum ($\nu_1 + \nu_2, \nu_1 + \nu_3, \dots$) and difference frequencies ($\nu_1 - \nu_2, \nu_1 - \nu_3, \dots$) allow the extension of the operating range of most of the lasers. Crystalline materials ADP, KDP, potassium pentaborate, β -barium borate, and lithium niobate are some of the crystals used for the frequency generation.

Metrology

The advantages of laser over other conventional sources in metrology are mainly because of the coherence, low divergence, and monochromatic nature of the laser light. He–Ne laser is most widely used because of its visible output, low power, and relatively low cost. Accurate measurement of distance by interference method is one area for which different types of commercial laser interferometers are available. They are used for precision measurement, calibration, testing, and for comparison with standards. Optical interference techniques using laser as a parallel monochromatic source have also been used for a number of less critical applications such as determination of diameter of very thin wires and small variations in surface smoothness. Laser scanning gauges to measure the roundness and diameter in different situations are also available.

11.13.3 Lasers in the Medical Field

The high energy density and directionality are the major characteristics that make laser an extremely useful device in surgery, ophthalmology, treatment of skin disorders, and so on.

The ability to obtain smaller focused beam size is an added advantage in some of the applications. The narrow intense beam can be used to destroy tissue in a localized area. It can be used to break up gallstones and kidney stones. Cancerous and pre-cancerous cells can be destroyed easily. During this process the heat that is generated in the process seals up capillaries and lymph vessels, which prevent the spread of the disease. For certain types of internal surgery, the laser beam is carried to the spot by an optical fiber.

A well accepted application of lasers in ophthalmology is for the welding of detached retina with the help of ruby or argon ion laser. Laser pulses of the order of 150 to 500 microseconds duration at levels below 1J are required and the patient can even be treated as an outpatient. As the power is low and duration is short, the ophthalmoscope can even be a hand held instrument. An argon ion laser is more suitable since the output of the ruby laser is not so effectively absorbed by blood vessels or haemoglobin. Such an arrangement for fusing retina offer the advantage of a smaller affected zone (0.15 to 1.5 mm), producing less damage to the healthy tissues. Cataract removal is possible with the help of low power He-Ne laser.

To view the internal parts of a human body, different types of flexible fibrescopes are available. Some of the commonly used fibrescopes are gastroscope, colonoscope, bronchoscope, laparoscope, etc. In all these fibrescopes, laser is the light source used.

11.13.4 Other Applications

Lasers offer a variety of applications and we have discussed some of them in the earlier sections. Some more applications are discussed below:

Holography

Another interesting application of laser light is in the field of holography. In holography, the images are formed by interference without lenses. The availability of coherent laser radiation made this possible. Holography helps in true reconstruction of three dimensional images, magnified or reduced in size or in full colour. Holography also makes possible the storage and retrieval of a large amount of information in a small volume (refer Section 12.4).

Laser induced nuclear fusion

The process of building up nuclei by the fusion of individual protons and neutrons or building large nuclei by combining small nuclei is called **nuclear fusion**. The enormous amount of energy released from the stars including the sun is due to the fusion that is taking place in them. Very high temperatures are required for fusion to occur and such devices are called **thermonuclear devices**. The sun and the stars are self sustaining thermonuclear reactors. However, one has to create such a situation for the reaction to take place on the earth. Thermonuclear reactors based on laser induced fusion offers an efficient method.

Optical communication

Optical communication is the transfer of information from one place to another using light waves as carrier. Since optical frequencies are of the order of 10^{15} Hz, the information carrying capacity of light wave is enormous. The high bandwidth available together with advantages such as

reliability, signal security, low transmission loss, electrical isolation, and immunity to interference make optical communication the most important mode of transmission of information from one place to another (for details refer Section 12.17).

Laser range finder

It is another useful device based on lasers. Optical range finder is often referred to as **LIDAR** (**LI**ght **D**etection **A**nd **R**anging) in analogy with radar. In range finders, the transit time of a reflected pulse of light is determined and the distance is evaluated. An accuracy of ± 5 m is possible at ranges of 8–10 km. More accurate methods are available using frequency modulated CW helium–neon laser. Methods exist for range finding and obstacle detection in fog, smoke and underwater. Large range finders are standard equipments for military surveillance. Optical radars are in use in aeroplanes, ambulances, police cars etc. Use of lasers for altimeters at low levels for aircraft is quite common. Based on Doppler effect, velocity measuring devices using low-power He–Ne laser are in operation. Laser Doppler techniques have been used for remote measurement of wind velocity and turbulence.

Military applications

Laser is used to mark targets for attack by artillery and guided missiles. In this, the laser sends a series of specific code of pulses of invisible light (IR). Special detecting systems are locked on to these pulses to guide the missiles to hit the target. Lasers are also used to build very high power devices that can send beams over very long distances at a very high speed. Another area is the use of laser for blinding enemy soldiers and their optical equipments.

Miscellaneous Applications

Listed below are some of the common applications of laser:

- (i) In scientific research as high intensity, monochromatic coherent light source
- (ii) As a barcode scanner in supermarkets and other places
- (iii) In laser printer
- (iv) In optical storage drives, CD players, CD rom drives, etc.
- (v) In photolithography
- (vi) In remote sensing.

SOLVED EXAMPLES

Example 11.1 Evaluate the ratio of spontaneous emission to stimulated emission at a temperature of 300 K for transitions that occur in (i) the visible region (ii) the microwave region.

Solution: (i) Taking the wavelength of visible region as $5500 \text{ \AA} = 5500 \times 10^{-10} \text{ m}$, we get

$$\frac{h\nu}{kT} = \frac{hc}{\lambda kT} = \frac{(6.626 \times 10^{-34} \text{ Js})(3 \times 10^8 \text{ m/s})}{(5500 \times 10^{-10} \text{ m})(1.38 \times 10^{-23} \text{ J/K}) 300 \text{ K}} = 87.3$$

$$\frac{\text{Spontaneous emission rate}}{\text{Stimulated emission rate}} = e^{\frac{hv}{kT}} - 1$$

$$\approx e^{87.3} \approx 8 \times 10^{37}$$

(ii) Taking the wavelength of microwave region as 1 cm, we get

$$\frac{hc}{\lambda kT} = \frac{(6.626 \times 10^{-34} \text{ Js})(3 \times 10^8 \text{ m/s})}{(10^{-2} \text{ m})(1.38 \times 10^{-23} \text{ J/K}) 300 \text{ K}} = 0.0048$$

$$\frac{\text{Spontaneous emission rate}}{\text{Stimulated emission rate}} = e^{0.0048} - 1 = 1.0048 - 1 = 0.0048$$

Example 11.2 Laser action occurs by stimulated emission from an excited state to a state of energy 30.5 eV. If the wavelength of laser light emitted is 690 nm, what is the energy of the excited state?

Solution: Energy of the laser light emitted is given by

$$= h\nu = \frac{hc}{\lambda} = \frac{(6.626 \times 10^{-34} \text{ Js})(3 \times 10^8 \text{ m/s})}{(690 \times 10^{-9} \text{ m})}$$

$$= 2.88 \times 10^{-19} \text{ J} = \frac{2.88 \times 10^{-19} \text{ J}}{1.6 \times 10^{-19} \text{ J/eV}} = 1.8 \text{ eV}$$

Energy of the excited state = 30.5 eV + 1.8 eV = 32.3 eV

Example 11.3 Determine the condition under which stimulated emission equals spontaneous emission.

Solution: Stimulated emission equals spontaneous emission when

$$e^{\frac{hv}{kT}} - 1 = 1 \quad \text{or} \quad e^{\frac{hv}{kT}} = 2$$

Taking logarithm on both sides, we get

$$\frac{hv}{kT} = \ln 2 = 0.693$$

Hence, the two types of emissions are equal when $hv = 0.693 kT$

$$\frac{v}{T} = \frac{0.693(1.38 \times 10^{-23} \text{ J/K})}{6.626 \times 10^{-34} \text{ Js}} = 1.44 \times 10^{10} \text{ K}^{-1} \text{ s}^{-1}$$

Example 11.4 Light from a 5 mW laser source of aperture 2 cm diameter and $\lambda = 500$ nm is focused by a lens of focal length 10 cm. Calculate (i) the area of the image (ii) the intensity at the focus.

Solution: Area of the spot at the focal plane is:

$$A = \frac{\pi\lambda^2 f^2}{a^2} = \frac{\pi(500 \times 10^{-9} \text{ m})^2 (0.15 \text{ m})^2}{(0.01 \text{ m})^2} = 1.77 \times 10^{-10} \text{ m}^2$$

Intensity at the focus $I = \frac{P}{A} = \frac{5 \times 10^{-3} \text{ W}}{1.77 \times 10^{-10} \text{ m}^2} = 2.82 \times 10^7 \text{ W/m}^2$

Example 11.5 The pulse width of a laser of wavelength 1064 nm is 25 ms. If the average power output per pulse is 0.8 W, (i) what is the energy released per pulse? (ii) How many photons does each pulse contain?

Solution: (i) Energy released per pulse = $P \times \Delta t$

$$= (0.8 \text{ W}) (25 \times 10^{-3} \text{ s}) = 20 \times 10^{-3} \text{ J}$$

$$\begin{aligned} \text{(ii) Number of photons in a pulse} &= \frac{\text{Energy}}{h\nu} = \frac{P \times \Delta t}{\frac{hc}{\lambda}} \\ &= \frac{(20 \times 10^{-3} \text{ J})(1064 \times 10^{-9} \text{ m})}{(6.626 \times 10^{-34} \text{ Js})(3 \times 10^8 \text{ m/s})} = 1.07 \times 10^{17} \end{aligned}$$

Example 11.6 Estimate the angular spread of a laser beam of wavelength 693 nm due to diffraction if the beam emerges through a 3 mm diameter mirror. How large would be the diameter of this beam when it strikes a satellite 300 km above the earth?

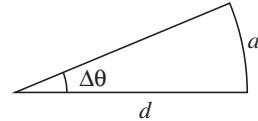
Solution: Diffraction effects limit the resolving power of an optical system. The minimum angle of resolution provided by a lens of diameter D is given by

$$\theta_{\min} = \frac{1.22\lambda}{D}$$

$$\begin{aligned} \text{Angular spread } \Delta\theta &= \frac{1.22\lambda}{D} = \frac{1.22(693 \times 10^{-9} \text{ m})}{(3 \times 10^{-3} \text{ m})} \\ &= 2.82 \times 10^{-4} \text{ rad} \end{aligned}$$

Diameter of the beam on the satellite is given as

$$\begin{aligned} a &= \Delta\theta \cdot d \\ &= (2.82 \times 10^{-4} \text{ rad})(300 \times 10^3 \text{ m}) \\ &= 84.6 \text{ m} \end{aligned}$$



REVIEW QUESTIONS

- 11.1** Explain spontaneous emission and stimulated emission of radiation.
- 11.2** Light from an ordinary light source is incoherent, whereas light from a laser is coherent. Why?
- 11.3** Explain the principle of laser.
- 11.4** What are Einstein's coefficients? State the relation between them.
- 11.5** Larger the energy difference between two states, more likely is spontaneous emission compared to stimulated emission. Why?
- 11.6** Why is laser action more difficult in the ultraviolet region of the electromagnetic spectrum.
- 11.7** Why is stimulated emission usually not observed?
- 11.8** What is population inversion? Why is it necessary for the operation of a laser?
- 11.9** Explain briefly the pumping methods used in lasers.
- 11.10** What are the basic components of a laser system? Explain the function of each component.
- 11.11** Explain the role of a resonant cavity in a laser system.
- 11.12** Write a note on the different resonant cavity configurations.
- 11.13** What is a metastable state? Explain its importance in lasers.
- 11.14** With an energy level diagram, explain the working of a three level laser.
- 11.15** With an energy level diagram, explain the working of a four level laser.
- 11.16** What are the advantages of a four level laser over a three level laser?
- 11.17** Explain why the lower laser level should not be the ground state.
- 11.18** How is population inversion achieved in a He–Ne laser?
- 11.19** Why is helium needed in a He–Ne laser? Why not just the neon?
- 11.20** Helium–Neon laser is a four level laser. Explain.
- 11.21** In a three level laser, the higher excited state is the metastable state. Would such an atom make a good laser?
- 11.22** Outline briefly the characteristics of laser light.
- 11.23** How would you demonstrate that laser light is coherent? What would be the result of the same experiment using an ordinary monochromatic source?
- 11.24** The coherence, energy density and directionality makes the laser a very useful device. Substantiate.
- 11.25** Distinguish between spatial coherence and temporal coherence.
- 11.26** What are Einstein's coefficients? Derive an expression for the ratio of the coefficients.
- 11.27** Explain the principle of laser. What are the essential requirements of a laser? Explain how these requirements are achieved.

- 11.28** With an energy level diagram, outline the working of a ruby laser. Explain briefly the important drawbacks of a three level laser.
- 11.29** Explain with an energy level diagram, the working of a He–Ne laser. Discuss some of the industrial applications of lasers.
- 11.30** Describe the principle and the working of a semiconductor laser. Write a note on some of its applications.
- 11.31** With an energy level diagram, outline the working of a CO₂ laser. Discuss some of the applications of CO₂ lasers.

PROBLEMS

1. The normal temperature of an optical source is 1000 K. Show that, in the optical region, the emission is predominantly due to spontaneous emission.
2. How many photons are emitted each second from a 2.5 mW He–Ne laser of wavelength 632.8 nm?
3. For the He–Ne laser ($\lambda = 632.8$ nm), estimate the broadening of the wavelength due to the uncertainty principle, assuming that the metastable state has a lifetime of 1 ms.
4. Light from a 2.5 mW laser source of aperture 1.8 cm diameter and $\lambda = 5000$ Å is focused by a lens of focal length 20 cm. Calculate the area of the image. Also calculate the intensity at the focus.
5. The output of laser has a frequency width of 700 Hz. If the frequency of the laser source is 10^{14} Hz, what is the monochromaticity of the output light?
6. A laser beam can be focused on an area equal to the square of its wavelength for a He–Ne laser $\lambda = 6328$ Å. If the laser radiates energy at the rate of 1 mW, find the intensity of the focused beam.

C H A P T E R

12

Holography and Fibre Optics

Holography, though discovered in 1948, reached its full potential only after the discovery of lasers. In conventional photography, only the intensity distribution can be recorded. But in holography, both the amplitude and phase of a wave can be recorded. This possibility makes holography have varied applications. Another equally interesting area in physics is fibre optics. Although the optical fibre was developed and placed in the market in the 1950's as light guides, its potential was also realized with the advent of lasers. These light guides enabled people to peep into otherwise inaccessible places. The optical fibre confines electromagnetic energy within its surface and guides the light in a direction parallel to its axis. Fibre optics has now evolved into a system of greater importance and use. We now discuss briefly some of the salient features of both holography and fibre optics.

12.1 HOLOGRAPHY—AN INTRODUCTION

Holography was discovered by Dennis Gabor in 1948, when he was attempting to increase the resolving power of electron microscopes. Holography is the process in which the amplitude and phase of a wave can be recorded using interferometric principle. In an ordinary photograph, the film records the intensity of reflected light from an object. When the photograph or transparency is viewed, a two dimensional image is obtained. In holography, a beam from a laser is split into two beams, a reference beam and an object beam. The object beam reflects or scatters light from the object to be photographed. The interference pattern produced by this light and the reference beam is photographed. Since the intensity at any point in the interference pattern also depends on the phase of the object wave, the resulting recording (**hologram**) contains information on the amplitude and phase of the object wave. When the hologram is

illuminated with the original reference wave, it reconstructs the original object wave. An observer looking through the hologram sees a perfect three dimensional image.

12.2 BASIC DESCRIPTION OF HOLOGRAPHY

12.2.1 Recording of Transmission Hologram

The physical set up for producing a transmission hologram is shown in Fig. 12.1. Light from the laser source is split into two components by a beam splitter. One of the components is directed towards the object and the other is directed to a suitable recording medium, usually a photographic emulsion. The wave that is directed to the recording medium is the **reference wave**. The beam that is directed to the object is reflected or scattered by the object. The scattered wave constitutes the **object wave**, which is also allowed to fall on the recording medium. Since the object wave and reference wave are mutually coherent, stable interference patterns are formed when the two meet at the recording medium.

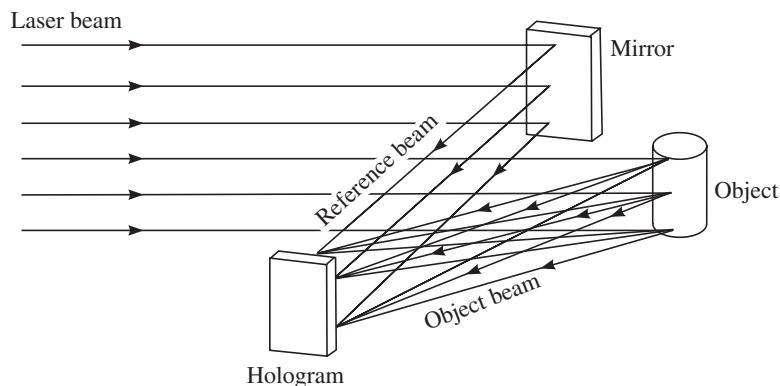


Fig. 12.1 Recording of the transmission hologram.

12.2.2 Reconstruction

In the reconstruction process, the hologram is illuminated by a beam of light similar to the reference wave used for recording the hologram. When the hologram is illuminated, light will only be transmitted through the clear portions. This results in a complex transmitted wave having three components (Fig. 12.2). Of these, one is the reconstruction wave itself. Another one exactly duplicates the original object wave, called the **reconstructed object wave**, which gives an exact replica of the original object in its true three-dimensional form. By placing a lens on the path of the reconstructed object wave one can get an image of the object on the screen. The third wave is the complex conjugate of the object wave, which gives a real image that can be photographed by placing a photographic plate at that position. Thus, holography is a two-step process by which images can be formed without lenses.

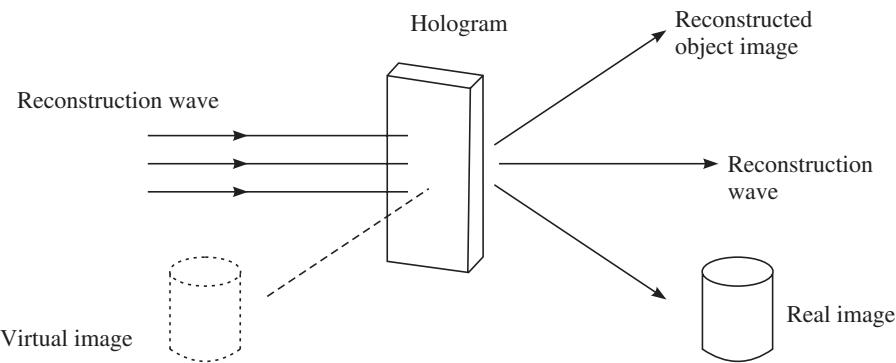


Fig. 12.2 Reconstruction of hologram the hologram diffracts the incident beam, producing real and virtual images.

12.3 TYPES OF HOLOGRAMS

12.3.1 Reflection Hologram

The reflection hologram is intended for viewing with reflected light. In a transmission hologram, the object beam and reference beam are incident on the same side of the recording film. In a reflection hologram, the object beam and reference beam are incident on opposite sides of the recording film. The important advantage of a reflection hologram is that laser light is not needed to view it. A point source of white light is sufficient. Moreover, it gives a less grainy appearing image than a transmission hologram.

12.3.2 In-line and Off-axis Holography

Based on the experimental arrangement, we have in-line holography (Gabor, 1948) and off-axis holography. In the in-line procedure, the reference wave is approximately parallel to the object wave and the path lengths of both the waves are almost equal. Consequently, the waves that form the virtual image and the real image travel along the same direction, which hinders the viewing of the virtual image. In the off-axis holography, the reference wave falls obliquely on the photographic plate. Here, the virtual image and the real image are well separated.

12.3.3 Diffuse Illumination of the Object

In a hologram recorded as detailed above a small region would have a one to one correspondence with a small region of the object. However, when the object is diffusely illuminated, each point on the object scatters light on the entire surface of the photographic plate. Hence, if the hologram is cut into small pieces, each piece would still reconstruct the whole object. However, the resolution of the image decreases as the size of the hologram decreases. Another advantage of such holograms is that the imperfections caused by dirt or scratches in the mirrors and beam splitters are smeared out over the whole plate, giving only a negligible effect on reconstruction.

12.3.4 Plane and Volume Holograms

Holograms for which photographic emulsion thickness is small are **plane holograms**. The recording medium is two dimensional. **Volume holograms** are those for which the thickness of the photographic emulsion is considerable. In these holograms, the information is stored throughout the volume of the emulsion. Hence the name volume hologram. Volume holograms can be viewed with ordinary white light. However, they must be made with laser light.

12.4 APPLICATIONS OF HOLOGRAPHY

Holograms exhibit interesting properties, and therefore, holography finds application in many diverse fields. We shall discuss some of the important ones in this section.

12.4.1 Holographic Interferometry

Holography is widely used in making precise interferometer measurements. Conventional interferometry is used to make measurements on highly polished surfaces of simple shapes. With holographic interferometry one can make measurements on three dimensional surfaces of arbitrary shapes and surface conditions. A major commercial application is its use in nondestructive testing. For example, objects undergoing pressure or heat stress in machines can be monitored holographically. The changing interference patterns reveal the faults in the parts. Different procedures are followed for the application of holographic interferometric techniques. In the double exposure interferometry, two exposures of the same emulsion are made with the same reference wave, one with the original object and the other with the object to be compared. The reconstructed object wave from the two interfere and produce the interference pattern, which can be used to compare the two objects. Usually, the two objects correspond to the same object under different conditions of strain.

12.4.2 Data Storage

One of the most exciting applications of holography is data storage. Holography records information in a distributed fashion, with each point on the hologram receiving light from all parts of the object. If part of the hologram is obscured or destroyed, it is still possible to recover all the data simply by looking through the remaining portion. Errors due to misfocussing and other defects in the recording material limit the use of microfilm for data storage. The data stored in a hologram can be pages of text. Many pages of data can be stored in one piece of material by utilizing the angular selectivity of thick holograms.

12.4.3 Holographic Optical Element

Holograms can be recorded to provide the same function as refractive optical elements such as lenses, prisms, and gratings. For example, consider a hologram recorded by the interference of a plane wave and a converging spherical wave. When the developed hologram is illuminated with a plane wave, it gives a converging spherical wave similar to that given by a lens. It is called a **Holographic Optical Element** (HOE). Several HOEs can be recorded on the same

piece of material. A HOE recorded in film can replace a short focal length lens with a large aperture, which is quite massive when fabricated in glass.

12.4.4 Holographic Microscopy

In contrast to a conventional high power microscope, a holographic microscope has an appreciable depth of field. The magnification (m) associated with reconstruction is given by

$$m = \frac{\text{Wavelength of light used for reconstruction}}{\text{Wavelength of light used for recording}}$$

To increase magnification, one has to do recording at short wavelength and reconstruction at a much longer wavelength. The non-availability of suitable coherent sources at very short wavelength (UV and X-rays) limits the magnification. Another drawback is the presence of wavefront aberrations introduced by the change between the recording wavelength and the reconstruction wavelength.

12.4.5 Pattern Recognition

In pattern recognition, the character to be read is matched with a labelled one called **reference**. It has become a routine practice in every discipline. In industry, it is used for the inspection of VLSI masks, wired boards, surface smoothness, etc. Finger print recognition requires the matching of a fingerprint in a labelled database. In biomedicine, it is used for a variety of purposes. Cell counting requires the recognition of different kinds of blood cells. In electrocardiography, the shape of the waveform helps the diagnosis. In military, target recognition is a frequently cited application.

In addition, holography is used for determining voids in layered objects including the inspection of composite aircraft components, multiple layer circuit boards, clutch plate spacing, nuclear fuel rods, antique paintings, etc. Studies for flow visualization in connection with air foils, plasmas, and combustion flames are quite common. Analysis of particle contamination of food and drug products are some of the other applications. Holograms on credit cards and other similar products are made with the help of holography in order to avoid forgery.

12.5 OPTICAL FIBRE

An optical fibre is a transparent rod, usually made of glass or clear plastic. It consists of a dielectric cylinder of refractive index n_1 , called the **core** of the fibre, surrounded by a solid dielectric called **cladding** having refractive index n_2 (where $n_2 < n_1$). The most widely used configurations of the fibre are the step index fibre and the graded or variable index fibre. In the step index fibre, the core is homogeneous with a constant refractive index n_1 and the cladding also has a constant refractive index n_2 . As there is an abrupt refractive index change at the core-cladding interface, the name is given as step index fibre. Figure 12.3(a) shows such a fibre along with the index profile. In the graded index fibre (GRIN fibre), the core refractive index varies radially, decreasing from the axis outward. In this type, the most commonly used is the one with a parabolic refractive index profile. Figure 12.3(b) illustrates a variable index fibre along with the index profile.

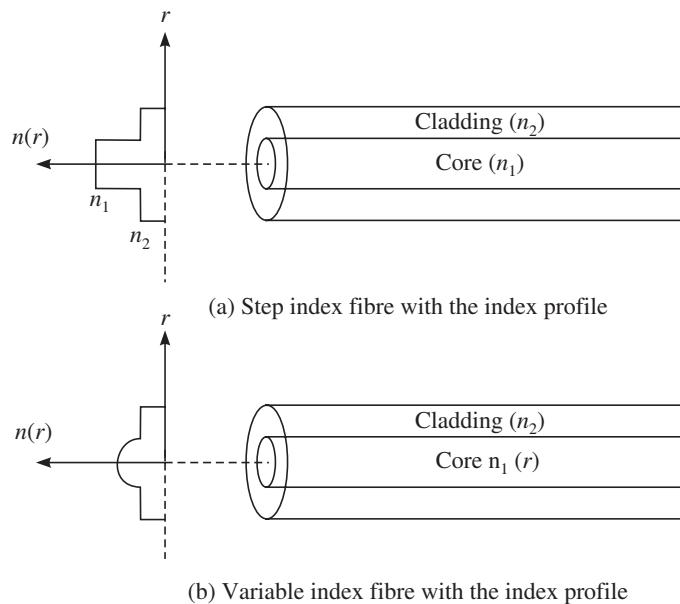


Fig. 12.3 Refractive index profiles of step index fibre and variable index fibre.

To protect the fibre and to give mechanical strength to it, an outer jacket is provided as shown in Fig. 12.4. Sufficient cushioning is also provided between the fibre and the outer jacket.

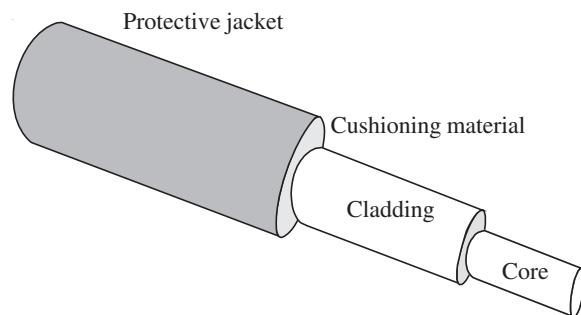


Fig. 12.4 Optical fibre with the outer jacket.

12.6 LIGHT PROPAGATION IN STEP INDEX FIBRE

In step index fibre, most of the light travels inside the core and is contained there by total internal reflection. Figure 12.5 illustrates the transmission of light in a step index fibre. At the air-core interface point A , refraction takes place and the rays continue at a smaller angle to the axis. They reach the core-cladding interface at points B_1 , B_2 , and B_3 . For the ray at B_1 , the angle of incidence is less than the critical angle α_c for the interface. Hence, it is refracted into the cladding, and finally it leaves the fibre or travels through the cladding. For the ray that reaches

B_2 , the angle of incidence is equal to the critical angle α_c . The angle θ_c is called the **critical propagation angle**. However, the ray that reaches B_3 is totally reflected into the core, as the angle of incidence is greater than the critical angle. That is, rays with propagation angles less than θ_c will be confined to the core of the fibre.

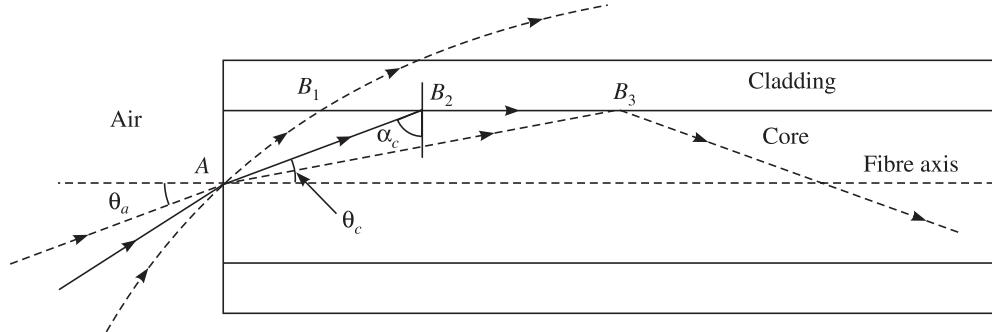


Fig. 12.5 Light propagation in a step index fibre.

12.7 NUMERICAL APERTURE

Consider a light ray that enters the fibre at an angle θ_a with respect to the fibre axis and strikes the core-cladding interface at an angle α_c (Fig. 12.6). By Snell's law, we have

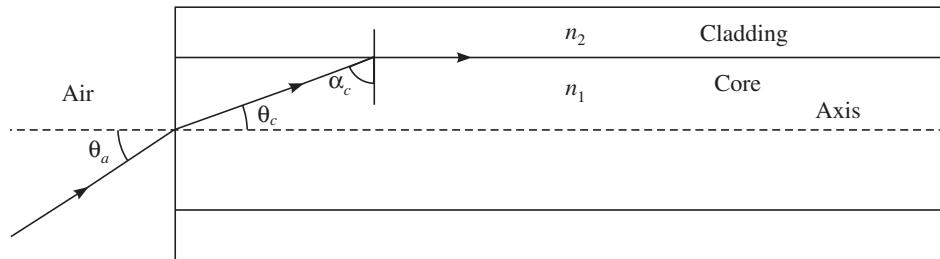


Fig. 12.6 Propagation of a ray in a step index fibre.

$$\sin \alpha_c = \frac{n_2}{n_1} \quad (12.1)$$

Since $\sin \alpha_c = \cos (90 - \alpha_c) = \cos \theta_c$, from Eq. (12.1), we get

$$\cos \theta_c = \frac{n_2}{n_1} \quad (12.2)$$

$$\sin \theta_c = \sqrt{1 - \cos^2 \theta_c} = \sqrt{1 - \left(\frac{n_2}{n_1}\right)^2} \quad (12.3)$$

If the launching is from air and n_0 is the refractive index of air, then

$$\frac{\sin \theta_a}{\sin \theta_c} = \frac{n_1}{n_0}$$

$$n_0 \sin \theta_a = n_1 \sin \theta_c = n_1 \sqrt{1 - \left(\frac{n_2}{n_1}\right)^2} = \sqrt{n_1^2 - n_2^2} \quad (12.4)$$

The term $n_0 \sin \theta_a$ is defined as the **numerical aperture** (N.A) of the fibre. Since $n_0 \approx 1$,

$$\text{N.A} = \sin \theta_a = \sqrt{n_1^2 - n_2^2} \quad (12.5)$$

Numerical aperture is a dimensionless quantity having values ranging from 0.14 to 0.5.

Another important parameter associated with light propagation in a fibre is the **relative refractive index difference** (Δ) defined by

$$\Delta = \frac{n_1 - n_2}{n_1} = \frac{(n_1^2 - n_2^2)}{n_1 (n_1 + n_2)} \approx \frac{(\text{N.A})^2}{2n_1^2} \quad (12.6)$$

since

$$n_1 + n_2 \approx 2n_1$$

12.8 ACCEPTANCE ANGLE AND ACCEPTANCE CONE

The rays that enter the fibre edge within the angle $2\theta_a$ (Fig. 12.6) will be accepted by the fibre. The angle $2\theta_a$ is called the **acceptance angle**. In three dimensions, it is an **acceptance cone** with semi-vertical angle θ_a . Acceptance cone is illustrated in Fig. 12.7. From Eq. (12.5), we get

$$\theta_a = \sin^{-1} \sqrt{n_1^2 - n_2^2} \quad (12.7)$$

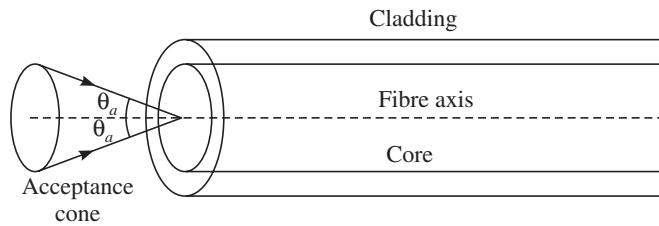


Fig. 12.7 Acceptance cone.

That is, light aimed at the fibre within this cone will be accepted and propagated to the far end. The larger the acceptance cone, easier the launching of light into the fibre. It may be noted from Eq. (12.5) that the N.A and the acceptance angle are independent of the fibre dimensions.

12.9 DIFFERENT MODES

As we have learnt, all rays with propagation angles less than θ_c will propagate through the fibre. Let us consider three such rays propagating at angles θ_1 , θ_2 , and θ_3 (Fig. 12.8). These

rays are called **modes of propagation**. The number of modes propagating in a fibre increase as θ_c increases. Since θ_c depends on n_1 and n_2 , the number of modes also depends on n_1 and n_2 . The total number of modes increases as the relative refractive index difference increases. The higher order modes are those with propagation angles close to the critical propagation angle θ_c , and the lower order modes are those with angles much lower than θ_c . On one hand, for lower order modes the fields are tightly concentrated near the centre of the fibre. On the other hand, for higher order modes the fields are distributed more toward the edge of the guide and tend to send light energy into the cladding. This energy is finally lost mainly at the fibre bends. The rays travelling at larger propagation angles (higher order modes) have to traverse a larger path length, and hence take a longer time as compared to those rays that propagate with lesser angles. Therefore, the higher order modes arrive later at the other end of the fibre. This leads to distortion of the signal called **intermodal distortion**. Although we have marked a mode by a ray, a mode actually means the nature of the electromagnetic field pattern inside the fibre.

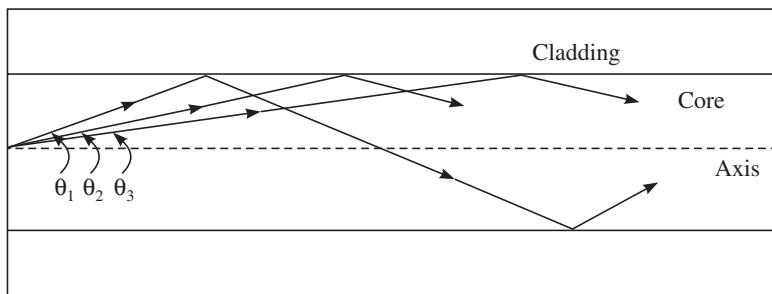


Fig. 12.8 Propagation modes in a step index fibre.

12.10 LIGHT PROPAGATION IN VARIABLE INDEX FIBRE

In a variable index fibre, the refractive index varies radially, decreasing from the axis outward. At the fibre axis, the refractive index is n_1 , at the cladding n_2 (where $n_1 > n_2$) and in between $n(r)$. As the ray goes from a region of higher refractive index to a region of lower refractive index, it is bent away from the normal. The process continues until the condition for total internal reflection is met, and the ray travels back towards the core axis, again being continuously refracted. This turning around may take place even before reaching the core-cladding interface. This continuous refraction followed by total internal reflection and again continuous refraction towards the centre is shown in Fig. 12.9. In a variable index fibre, although rays making larger angles with the axis traverse a longer path length, they do so in a region of lower refractive index, and this results to a higher speed propagation. Consequently, all rays travelling through such a fibre will have almost the same optical path length and reach the other end almost at the same time. Hence, variable index fibres are more suitable for high bandwidth applications.

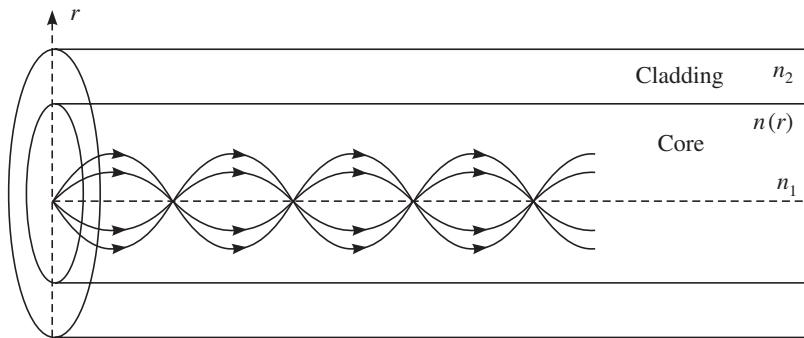


Fig. 12.9 Light transmission in a multimode graded index fibre.

12.11 SINGLE AND MULTIMODE FIBRES

As the name implies, a single mode fibre sustains only one mode of propagation, whereas a multimode fibre contains multiple modes.

12.11.1 Salient Features of Single Mode Fibres

- (i) Typical core diameter of a single mode step index fibre is 8-12 μm , and the cladding thickness is about 125 μm . In a graded index single mode fibre, the core diameter is slightly larger than that in a step index fibre. This makes handling, connecting, and coupling a graded index fibre slightly easier than a step index fibre.
- (ii) Fibre losses will be less.
- (iii) The cladding is relatively thick in a single mode fibre, as part of the light propagates in the cladding. By making the cladding thick, the field at the cladding air boundary can be minimized.
- (iv) Both N.A and Δ are very small. Low N.A means a very small acceptance angle. That is, the incident ray must be nearly perpendicular to the fibre edge.
- (v) As only one mode exists in the fibre, no intermodal dispersion exists. This makes single mode fibres suitable for use with high data rates.
- (vi) Manufacturing a single mode fibre is more expensive and difficult. To launch light energy into the fibre, costly laser diodes are needed. If the high data rate is taken into consideration, it is cost effective.
- (vii) Single mode fibres are becoming more popular for other specialized applications.

12.11.2 Salient Features of Multimode Fibres

- (i) In a multimode step index fibre, the core diameter is of the order of 50 - 200 μm , with a cladding thickness of 125 - 400 μm . The core diameter of a graded index fibre is 50 - 100 μm , with a cladding thickness of 125 - 140 μm .

- (ii) As N.A and Δ are larger for a multimode fibre, its acceptance angle is much larger than that for the single mode. Hence, launching of a light signal into the fibre is much easier.
- (iii) Another advantage is that an LED, which is comparatively less expensive, can be used to launch the light.
- (iv) Handling, connecting, and coupling of multimode fibres are easier than those of compared to single mode fibres.
- (v) The major disadvantage is intermodal dispersion, which limits data transmission rates.

12.12 V-NUMBER OF A FIBRE

The *V*-number of a fibre is defined by

$$V = \frac{2\pi a}{\lambda} (n_1^2 - n_2^2)^{\frac{1}{2}} = \frac{2\pi a(N.A)}{\lambda} \quad (12.8)$$

where a is the fibre core radius and λ is the operating wavelength. The *V*-number is dimensionless, and is also referred to as **normalized frequency**. The parameter V is related to the number of modes a fibre can support. Detailed calculations give the number of modes M supported in a multimode step index fibre as

$$M_{SI} = \frac{1}{2} \frac{(2\pi a)^2}{\lambda^2} (n_1^2 - n_2^2) = \frac{V^2}{2} \quad (12.9)$$

If the two possible polarizations are taken into account, the number of modes gets doubled.

$$M_{SI} = V^2 \quad (12.10)$$

The number of modes in a graded index fibre is about half of that in a step index fibre.

$$M_{GI} = \frac{V^2}{4} \quad (12.11)$$

Taking the two polarizations into account, we get

$$M_{GI} = \frac{V^2}{2} \quad (12.12)$$

12.13 FIBRE LOSSES AND TRANSMISSION WINDOWS

The wavelength range that is often used for communication is 0.7 to 1.7 μm . Hence, we need to consider only the fibre losses in this range. The major contributors to the loss of light in a fibre during transmission are (i) material loss, (ii) light scattering loss, and (iii) loss due to geometric effect. The material impurities and the OH^- ions present in the fibre are the major contributors for the material loss in fibres. The most important impurity is the trapped OH^- ions and the losses due to them occur around 0.95, 1.23 and 1.39 μm (Fig. 12.10).

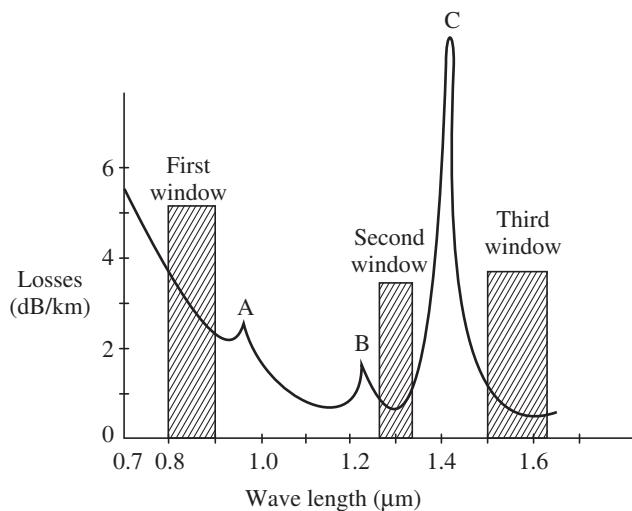


Fig. 12.10 Fibre losses over 0.7 to 1.6 μm region. A, B, C – OH^- absorption peaks.

Geometric effects include losses due to macro and micro bends. Combining all of the loss phenomena, we find only a small loss in the following regions for glass fibres:

- 0.80 to 0.90 μm region, called the **first window**
- 1.26 to 1.34 μm region, the **second window**
- 1.50 to 1.63 μm region, the **third window**

For calculating the signal attenuation, a parameter α , called **fibre loss** or **fibre attenuation** is generally used. If P_i is the power input to the fibre and P_0 is the fibre output at a distance z

$$\alpha = \frac{10}{z} \log \frac{P_i}{P_0} \quad (12.13)$$

Usually, a low loss fibre may have a 1 dB/km average loss at 900 nm. These days fibres having losses as low as 0.1 dB/km are very common.

12.14 OPTICAL SOURCES FOR FIBRE OPTICS

The selection of an optical source can be done by taking the following aspects into account:

- (i) Dimensional compatibility with the fibre core.
- (ii) Fibre's attenuation as a function of wavelength.
- (iii) Must be highly coherent and monochromatic.
- (iv) The purpose for which the system is used.
- (v) Should easily be coupled to the fibre.
- (vi) Must have a faster response time.
- (vii) Must be highly directional.

The light sources used for fibre optic applications are the light emitting diodes (LEDs) and diode lasers, also called injection laser diodes. These are suitable for fibre applications because

- (i) They have the required power.
- (ii) They have a high efficiency.
- (iii) Dimensional characteristics are compatible with those of the optical fibre.
- (iv) Optical power output can be directly modulated by varying the input current to the device.

The optical output is strongly temperature dependent. Hence, cooling must be used to maintain the laser at a constant temperature. The major differences between LED and diode laser are:

- (i) The output from the LED is incoherent, whereas that from a diode laser is coherent.
- (ii) The light from an LED has a broad spectral width and beam divergence, and therefore the information carrying capacity of a system is much less.
- (iii) The light from a diode laser has both spatial and temporal coherence, and so is highly monochromatic.
- (iv) Diode laser is highly directional.
- (v) Coupling the LED to the fibre is more difficult and the amount of power it can launch into the fibre is relatively small.
- (vi) LED is less expensive compared to laser diodes.

12.15 OPTICAL COMMUNICATION

The most interesting application of fibre optics is optical communication, which is the transfer of information from one place to another using light wave as carrier. To transmit one speech signal, a bandwidth of about 4 kHz is required. Similarly, bandwidths of about 15-20 kHz and 6 MHz are required to transmit music and television signals, respectively. The microwave channel has the frequency range 10^9 to 5×10^9 Hz. In this channel we can send about 10^6 speech signals or 10^5 music signals or 700 TV signals. The visible region corresponds to the frequency range 4.2×10^{14} to 7.5×10^{14} Hz. The number of signals that can be transmitted in this range is $10^{14} \text{ MHz}/10^6 \text{ MHz} = 10^8$ TV signals. That is, the information carrying capacity of the optical region is enormous. This process has led to the birth of optical communications in which optical fibre is the medium that guides the light waves.

12.15.1 Communication System

A schematic block diagram of an optical communication system is shown in Fig. 12.11. Basically, it consists of three blocks—the transmitter, the optical fibre, and the receiver. The modulation of the optical carrier wave may be done using either an analog or a digital information signal. Although analog modulation is simpler to implement, often it is not preferred as it is less efficient. It requires a higher signal to noise ratio at the receiver and perfect linearity of the optical source. Due to these reasons, analog optical fibre communication links are often restricted to shorter distances and lower bandwidths.

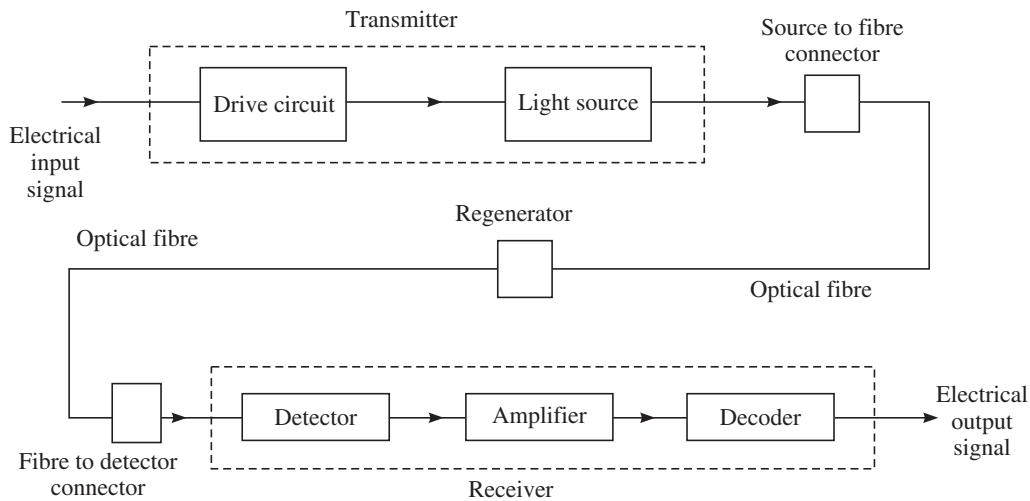


Fig. 12.11 Block diagram of an optical communication system.

In any system, there is a maximum allowed distance between the transmitter and receiver, beyond which the system efficiency decreases considerably. Even with the highest intensity light sources and lowest loss fibres, the light waves finally become so weak that they must be regenerated. At this point a regenerator is placed in the circuit. It consists of a light receiver, a photon to electron conversion unit, electrical amplification, pulse shaping, and an electron to photon conversion unit. Together they rebuild the pulses to their original level and send them down the transmission line. While designing a communication system, one keeps in mind better transmission fidelity, increased data rate, and maximum distance between regenerator stations. The information transfer capacity is further increased by multiplexing several signals on to a single fibre. The method of sending more than one signal simultaneously along the same channel is called **multiplexing**.

12.15.2 Advantages of Optical Communication

The purpose of any communication channel is to send information over a required distance. One can assess its performance by the amount of information that can be transmitted and how effectively it is done. In the light of these, we list below the merits of optical fibre communication over the conventional methods (radio and microwave):

- (i) **Higher information carrying capacity (bandwidth):** As explained, the information carrying capacity of light waves is extremely high, as the bandwidth available is extremely large.
- (ii) **Small size and weight:** Fibres are very thin. Their diameter is less than that of human hair. Hence, even with protective coating they are smaller and much lighter than copper cables. Therefore, they occupy only very little space in the already crowded ducts in cities.
- (iii) **Immunity to interference:** As optical fibre is a dielectric waveguide, it is free from electromagnetic pulses, etc. It can be treaded through electrically noisy environment.

Fibre cable requires no shielding from electromagnetic interference. It is free from lightening strikes if used overhead.

- (iv) **Negligible cross talk:** Optical interference between fibres is negligible, as fields are confined to the fibres. Hence, cross talk is negligible, even when a number of fibres are cabled together.
- (v) **Low transmission loss:** The development of optical fibres with losses as low as 0.2 dB/km or less allows the implementation of communication links with extremely large repeater spacing.
- (vi) **Ruggedness and flexibility:** With protective coatings and proper cabling, manufacture of optical fibres of high tensile strength is possible. In terms of storage, transportation, handling, and installation, optical cables are generally superior to copper cables.
- (vii) **Signal security:** Light from optical fibres do not radiate significantly, and hence it is difficult to tap information during transmission.
- (viii) **System reliability:** The reliability of optical components is 20 to 30 years. Hence, system reliability is much better than that with conventional electrical conductor systems.
- (ix) **Low cost:** The cost of semiconductor lasers, photodiodes, connectors, and couplers are on the higher side. The number of repeater stations required is less. Taking all the factors such as the information carrying capacity, system reliability, predicted lifetime, and security, the overall system cost is very low.

Owing to these advantages, light wave communication is fast becoming an important mode of transmission of information from one place to another.

12.16 LIGHT PIPE

A major area of application of optical fibre is either to illuminate inaccessible places or to conduct light out of such places. In such a system, an optical fibre is used as a light pipe. The most interesting example in this series is the flexible fibrescope. Fibrescopes are extensively used in medicine and engineering to peep into the interior of a human body or the interior of a jet engine. In these, one set of fibres is used to carry light into the area to be examined and another set is used to carry the image back to the observer. Different types of flexible fibrescopes are available. Following are some of the ones used to view the internal parts of the human body:

- (i) Internal parts of the stomach—Gastroscope
- (ii) Large intestine—Colonoscope
- (iii) Respiratory tract and parts of the lungs—Bronchoscope
- (iv) Internal organs of the abdomen—Laparoscope.

12.17 FIBRE OPTIC SENSORS

Fibre optic sensors are devices in which laser light guided by the fibre undergoes a modulation in response to an external perturbation. They are now extensively used in the measurement of

displacement, temperature, pressure, presence or absence of an object, etc. Sensors can be classified according to the optical property of the guided light that is modulated. Consequently, we have the following types of sensors:

- (i) **Intensity modulated sensors:** In this type of sensors, the intensity of the guided light is modulated by change in force, temperature, etc.
- (ii) **Phase modulated sensors:** These are based on the change of phase of the light wave guided by the fibre.
- (iii) **Polarization modulated sensors:** These sensors are based on the change in polarization of the light guided by the fibre.
- (iv) **Wavelength modulated sensors:** In these sensors, the modulation of the wavelength of the light wave that is guided by the fibre by some external process is studied to measure the details of the external process.

12.17.1 Intensity Modulated Sensor

Perhaps the most popular sensors are those that are based on variations of light intensity. Light from a source is split into two parts by a beam splitter, one of which serves as a reference and the other is allowed through a fibre without jacket, which is placed between two corrugated blocks (refer Fig. 12.12). Force applied on the top block presses the fibre, which introduces microbend losses in it. The changes in intensity due to these losses are measured with respect to the reference signal. The force can easily be evaluated, as the change in intensity is related to the force.

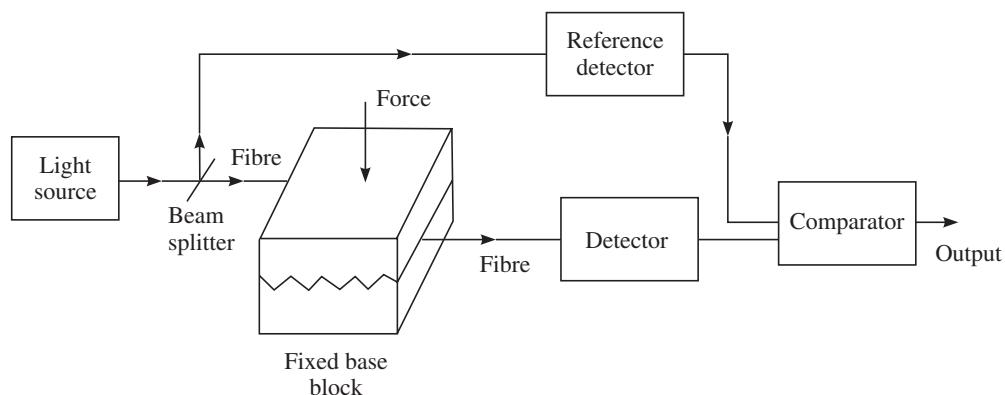


Fig. 12.12 Force sensor using microbend losses.

12.17.2 Phase Sensor

As an example, consider a temperature sensor based on phase variation, which is illustrated in Fig. 12.13. Light from a laser is split into two parts by a beam splitter; one part is allowed through the sensor element of the fibre and the other through the reference fibre. As the sensor element is heated, the fibre elongates, thereby increasing the path length for light. This changes the phase

of the light that reaches the detector. The phase comparator measures the difference in phase, which is related to the temperature. The light source must be highly coherent.

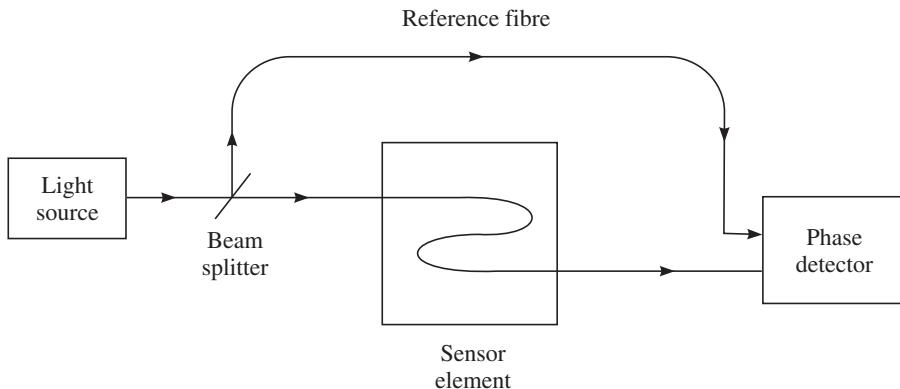


Fig. 12.13 Temperature sensor using phase variations.

12.17.3 Sensors in Health Care

With the development of new devices for use in medicine and health care, fibre optic sensors have been developed such as non-invasive, mini-invasive and invasive sensors. Non-invasive sensors are those that are simply placed in contact with skin, mini-invasive are the ones introduced inside the natural cavities of the human body, invasive sensors are inserted into organs though small incisions or through blood vessels. Since uncabled fibres are very small and flexible, they can be inserted inside very thin catheters and hypodermic needles, thereby allowing highly localized and minimally invasive monitoring. Absence of cross talk between neighbouring fibres allows several sensors to be grouped in a single catheter.

Analysis of blood gases is a normal diagnostic procedure in hospitals. Current monitoring techniques involve blood sampling and in vitro diagnosis. Fibre optic sensors allow continuous monitoring by inserting the fibres into the vessels, thus allowing in vivo diagnosis and immediate decisions. The success of radiotherapy relies on online monitoring of the dose to which a tumour and its adjacent tissues are exposed. Conventional dosimeters only provide offline monitoring since they determine radiation exposure after the completion of the exposure. Fibre optic sensors are now available for continuous monitoring of radiation dosage in both invasive and non-invasive applications. Fibre optic thermometers are used in mapping the heat distribution in tissues during cancer phototherapy, and patient monitoring during magnetic resonance imaging. A major ophthalmological application is the early recognition of cataract onset. A fibre optic sensor developed for dermatological applications helps in quantifying erythema, an acute inflammatory skin disease. Summing up, fibre optic sensors find extensive application in different areas of health care and the day is probably not far off in developing multipurpose portable equipment using fibre optic sensors.

SOLVED EXAMPLES

Example 12.1 For a step index fibre having core index 1.43 and cladding index 1.4, calculate (i) critical angle α_c (ii) critical propagation angle θ_c (iii) numerical aperture.

$$\text{Solution:} \quad (i) \sin \alpha_c = \frac{n_2}{n_1} = \frac{1.4}{1.43} = 0.979; \quad \alpha_c = 78.24^\circ$$

$$(ii) \cos \theta_c = \frac{n_2}{n_1} = \frac{1.4}{1.43} = 0.979; \quad \theta_c = 11.76^\circ$$

$$(iii) N.A = \sqrt{n_1^2 - n_2^2} = \sqrt{1.43^2 - 1.4^2} = 0.29$$

Example 12.2 The refractive index of the core and cladding of a fibre are 1.45 and 1.4, respectively. Find the numerical aperture, acceptance angle, and Δ .

$$\text{Solution:} \quad N.A = \sqrt{1.45^2 - 1.4^2} = 0.377$$

$$\sin \theta_a = \sqrt{1.45^2 - 1.4^2} = 0.377; \quad \theta_a = 22.15^\circ$$

$$\text{Acceptance angle } 2\theta_a = 44.3^\circ$$

$$\Delta = \frac{n_1 - n_2}{n_1} = \frac{1.45 - 1.4}{1.45} = 0.034$$

Example 12.3 A step index fibre has a core of refractive index 1.55 and a cladding of 1.53. If the signal is launched from a medium of refractive index 1.3, (i) what is the numerical aperture? (ii) what is the acceptance angle?

$$\text{Solution:} \quad (i) N.A = n_0 \sin \theta_a = \sqrt{n_1^2 - n_2^2} = \sqrt{1.55^2 - 1.53^2} = 0.248$$

$$(ii) n_0 \sin \theta_a = 0.248$$

$$\sin \theta_a = \frac{0.248}{1.3} = 0.1908; \quad \theta_a = 11^\circ$$

$$\text{Acceptance angle } 2\theta_a = 22^\circ$$

Example 12.4 A step index fibre having a core index of n_1 and cladding index n_2 is used in a communication link. Prove that the numerical aperture is approximately equal to $n_1(2\Delta)^{\frac{1}{2}}$ where $\Delta = \frac{(n_1 - n_2)}{n_1}$.

Solution: Numerical aperture $= (n_1^2 - n_2^2)^{\frac{1}{2}} = (n_1 + n_2)^{\frac{1}{2}} \left(\frac{n_1 - n_2}{n_1} \right)^{\frac{1}{2}} n_1^{\frac{1}{2}}$

$$\equiv (2n_1)^{\frac{1}{2}} \Delta^{\frac{1}{2}} n_1^{\frac{1}{2}} = n_1 (2\Delta)^{\frac{1}{2}}$$

Since

$$(n_1 + n_2)^{\frac{1}{2}} \equiv (2n_1)^{\frac{1}{2}}$$

Example 12.5 A communication system uses a 10 km fibre, which has a fibre loss of 2.0 dB/km. Find the output power if the input power is 500 μ W.

Solution: The power loss α (db/km) $= \frac{10}{z} \log \frac{P_i}{P_0}$

$$2 = \frac{10}{10} \log \left(\frac{500 \mu \text{W}}{P_0} \right) \quad \text{or} \quad \log 100 = \log \left(\frac{500 \mu \text{W}}{P_0} \right)$$

$$\frac{500 \mu \text{W}}{P_0} = 100 \quad \text{or} \quad P_0 = 5 \mu \text{W}$$

REVIEW QUESTIONS

- 12.1 Explain how a hologram is recorded and how is it reconstructed.
- 12.2 Distinguish between photography and holography.
- 12.3 Distinguish between transmission and reflection holograms.
- 12.4 Explain the concept of diffuse illumination of the object in holography. What are its advantages?
- 12.5 How does holographic interferometry differ from the conventional one?
- 12.6 Write a note on holographic data storage.
- 12.7 What is a holographic optical element? What are its advantages?
- 12.8 What is a step index fibre? How is light propagated in a step index fibre?
- 12.9 What is a variable index fibre? How is light propagated in it?
- 12.10 What is acceptance angle? How does it depend on fibre dimensions?
- 12.11 Why are variable index fibres preferred for high bandwidth applications?
- 12.12 Higher order modes tend to send light energy into the cladding. Why?
- 12.13 Define numerical aperture. How is it related to the acceptance angle?
- 12.14 The information carrying capacity of electromagnetic radiation in the optical region is enormous. Why?

- 12.15** Compare the performance of laser diode and LED as light source for fibre optic applications.
- 12.16** Outline the advantages of light wave communication over the conventional ones.
- 12.17** Write a brief note on fibre optic sensors.
- 12.18** What is a fibrescope? Mention some of the fibrescopes used in health care.
- 12.19** What is a hologram? Explain with suitable diagrams how a hologram is recorded and then reconstructed.
- 12.20** Explain how light is propagated in a step index fibre. Define numerical aperture and acceptance angle. Derive expressions for the same.
- 12.21** Explain how light is propagated in a variable index fibre. Discuss the salient features of single and multimode transmission.
- 12.22** Define numerical aperture, relative refractive index difference, and V-number of an optical fibre. Derive expressions for all the three.
- 12.23** Explain with a block diagram an optical communication system and discuss some of its advantages. Write a note on the optical sources used for optical communication.
- 12.24** Discuss the different types of losses in optical fibres. Explain the use of optical fibres as sensors.

PROBLEMS

1. In a step index fibre, the refractive index of the core and the cladding are 1.43 and 1.4, respectively. Find the numerical aperture and the acceptance angle.
2. A step index fibre has a core refractive index of 1.44 and cladding refractive index 1.41. Find (i) the numerical aperture, (ii) the acceptance angle, and (iii) the relative refractive index difference.
3. Find the numerical aperture and acceptance angle of a fibre of core index 1.4 and $\Delta = 0.02$.
4. A step index fibre has a core of index 1.55 and a cladding of index 1.53. (i) What is its numerical aperture? (ii) What is its acceptance angle?
5. A fibre has a core index of 1.4 and relative refractive index difference of 0.02. Find the refractive index of the cladding.
6. A variable index fibre has a refractive index of 1.5 along the core axis and a cladding refractive index of 1.46. What is its numerical aperture?
7. A fibre of 100 m length has a power input $P_i = 10 \mu\text{W}$ and power output $P_0 = 8.8 \mu\text{W}$. Find the power loss in dB/km.
8. A 30 km long optical fibre having a fibre loss of 0.8 dB/km is used in a communication link. If the input power is 200 μW , what would be the output power?
9. A communication system uses a 10 km fibre having a fibre loss of 2.5 dB/km. Find the input power if the output power is 1.2 μW .

C H A P T E R

13

Dielectric Properties of Materials

Dielectrics are insulators having interesting electrical properties, since an electric field can polarize the material to create electric dipoles. In this chapter we shall study how a dielectric behaves in an applied electric field. The relation between the applied field and the internal field within the dielectric will also be investigated.

13.1 REVIEW OF BASIC IDEAS

13.1.1 Electric Dipole

An electric dipole is composed of two equal and opposite charges q and $-q$ separated by a distance d . The moment of this dipole is defined as

$$p = qd \quad (13.1)$$

Electric dipole is a vector quantity, its direction being from the negative to the positive charge.

13.1.2 Polarizability

A molecule contains an equal number of positive and negative charges. Under the influence of an external electric field \mathbf{E} , the positive charges are displaced in one direction and the negative charges in the opposite direction, the molecule is said to be polarized and the phenomenon is called **polarization** forming an electric dipole. If d is the separation between the centres of the positive and negative charges, the induced dipole moment μ is given by

$$\mu = \alpha \mathbf{E} \quad (13.2)$$

where α is called the **polarizability** of the molecule.

13.1.3 Polarization Vector

The induced dipole moment per unit volume of the material is called the **polarization vector** \mathbf{P} . If N is the number of molecules per unit volume, the polarization vector is given by

$$\mathbf{P} = N\mu = \alpha N\mathbf{E} \quad (13.3)$$

13.1.4 Dielectric Constant

When a dielectric is introduced between the plates of a parallel plate capacitor, the value of its capacitance C increases by a factor ϵ , where ϵ is called the **dielectric constant or permittivity** of the medium.

$$C = \frac{\epsilon A}{d} \quad (13.4)$$

where A is the cross sectional area of the plates and d is the separation between the plates. A dielectric is also characterized by its **relative dielectric constant** or **relative permittivity** ϵ_r . It is defined as the ratio between the permittivity of the medium ϵ to the permittivity or dielectric constant of free space (vacuum) ϵ_0 . That is,

$$\epsilon_r = \frac{\epsilon}{\epsilon_0} \quad (13.5)$$

The dielectric constant of free space $\epsilon_0 = 8.854 \times 10^{-12}$ F/m. From Eqs. (13.4) and (13.5), we get

$$C = \frac{\epsilon_0 \epsilon_r A}{d} \quad (13.6)$$

13.1.5 Dielectric Susceptibility

When a dielectric is placed in an external electric field \mathbf{E} , the polarization \mathbf{P} is always proportional to the electric field. That is,

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} \quad (13.7)$$

where χ is the **dielectric susceptibility** of the dielectric medium. In general, it is a tensor. The factor χ characterizes the ease with which a dielectric material can be influenced by an external field. From Eq. (2.17), the electric displacement density vector \mathbf{D} for an isotropic medium is given by

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \epsilon_r \mathbf{E} \quad (13.8)$$

In terms of polarization vector \mathbf{P} , it is given by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (13.9)$$

From Eqs. (13.8) and (13.9), we have

$$\mathbf{P} = \epsilon_0 (\epsilon_r - 1) \mathbf{E} \quad (13.10)$$

Combining Eqs. (13.7) and (13.10), we obtain

$$\epsilon_r = 1 + \chi \quad (13.11)$$

Another useful relation can be obtained from Eqs. (13.3) and (13.10) as follows:

$$\epsilon_r - 1 = \frac{N\alpha}{\epsilon_0} \quad (13.11a)$$

13.2 DIELECTRIC POLARIZATION

As mentioned, the characteristic that all dielectric materials have in common is the displacement in the relative positions of the positive and negative charges, when an external electric field is applied to it. As a result, microscopic electric dipoles are formed, leading to a polarization of the medium. We shall now examine the physical processes that give rise to the polarization. The major contributors to the total polarization are of four types—electronic polarization, ionic polarization, orientational polarization, and space charge polarization. The amount of polarization by each type depends on the nature of the dielectric and the frequency of the field.

13.2.1 Electronic Polarization

The dielectrics have no free electrons in it. Hence, electrons are bound tightly to their parent atoms. As mentioned, in an electric field there is a net displacement of net negative charge relative to the net positive charge, giving rise to electric dipoles. The alignment of such atomic dipoles with respect to the direction of the electric field causes polarization of the material, called **electronic polarization**.

Let us consider a dielectric of rare gas atoms of an element in a static electric field \mathbf{E} applied in the positive x direction. In the gaseous state, the distance between atoms is large as compared to its dimensions. Hence, the interaction between atoms can be neglected. When $\mathbf{E} = 0$, let us assume that the nucleus of charge Ze is surrounded by an electron cloud of charge $-Ze$, distributed uniformly in a sphere of radius R .

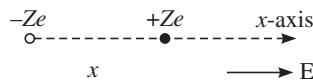


Fig. 13.1 Schematic representation of an atomic dipole in the presence of electric field \mathbf{E} .

The volume density of electronic charge can be written as

$$\rho = \frac{-Ze}{\left(\frac{4}{3}\right)\pi R^3} = \frac{-3Ze}{4\pi R^3} \quad (13.12)$$

When $\mathbf{E} \neq 0$, the nucleus and the electrons experience Lorentz forces $Ze\mathbf{E}$ (magnitude) in opposite directions, resulting in a displacement x between the nucleus and the centroid of the electron charge distribution (Fig. 13.1). When they are separated, a Coulomb attraction develops

between them, which opposes the displacement. Hence, at equilibrium, the Lorentz force must equal the Coulomb force.

$$\begin{aligned} \text{Coulomb force} &= \frac{Ze \times \text{Charge in the sphere of radius } x}{4\pi\epsilon_0 x^2} \\ &= \frac{Ze \times \left(\frac{4}{3}\right)\pi x^3 \rho}{4\pi\epsilon_0 x^2} \end{aligned} \quad (13.13)$$

Substituting the value of ρ , we get

$$\begin{aligned} \text{Coulomb force} &= -\frac{Z^2 e^2 x}{4\pi\epsilon_0 R^3} \end{aligned} \quad (13.14)$$

Equating the two forces, we get

$$\begin{aligned} -Ze E &= -\frac{Z^2 e^2 x}{4\pi\epsilon_0 R^3} \\ E &= \frac{Ze}{4\pi\epsilon_0} \frac{x}{R^3} \end{aligned} \quad (13.15)$$

$$x = \frac{4\pi\epsilon_0 R^3 E}{Ze} \quad (13.16)$$

That is, the displacement x is proportional to the applied electric field E . Induced electric dipole moment,

$$\begin{aligned} \mu_e &= Zex = 4\pi\epsilon_0 R^3 E \\ &= \alpha_e E \end{aligned} \quad (13.17)$$

where,

$$\alpha_e = 4\pi\epsilon_0 R^3 \quad (13.18)$$

is called the **electronic polarizability**. Electronic polarization is defined as electronic dipole moment per unit volume. If N is the number of atoms per m^3 , electronic polarization

$$P_e = N\alpha_e E \quad (13.19)$$

13.2.2 Ionic Polarization

Ionic polarization occurs due to the displacement of charged ions relative to other ions and is exhibited by crystalline dielectrics such as NaCl and CsCl. In ionic crystals, the positive and negative ions are bonded by ionic bond. When an electric field is applied along the x -axis, ions are displaced in opposite directions until ionic bonding stops the process. When the field is reversed, ions move closer and again the dipole moment undergoes a change. The induced dipole moment which is proportional to the applied field is given as by

$$\mu_i = \alpha_i E \quad (13.20)$$

where α_i is the ionic polarizability. The ionic polarization is:

$$\mathbf{P}_i = N\alpha_i \mathbf{E} \quad (13.21)$$

13.2.3 Orientation Polarization

This type of polarization occurs only in polar substances. In such molecules, the centres of positive and negative charges do not coincide and each pair of charges acts as a dipole. Under normal circumstances, these dipoles are oriented in a random way, as shown in Fig. 13.2(a). When an electric field is applied to such a material, these dipoles are aligned, to some extent, in the field direction. The material is then said to be polarized and the phenomenon is known as **orientational polarization**. This is illustrated in Fig. 13.2(b).

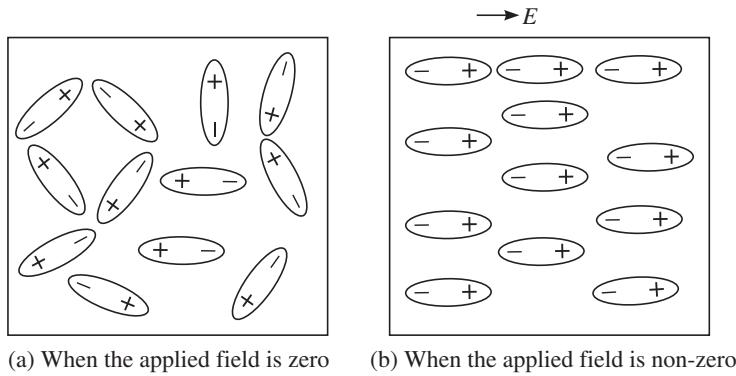


Fig. 13.2 Orientational polarization.

Orientation polarization at not-too-low temperatures and not-too-low fields can be written as

$$\mathbf{P}_o = N\alpha_o \mathbf{E} \quad (13.22)$$

where α_o is the orientation polarizability defined by

$$\alpha_o = \frac{\mu^2}{3kT} \quad (13.23)$$

where μ is the dipole moment of the molecule, k is the Boltzmann constant, and T is the temperature. It is evident from Eq. (13.23) that the orientation polarization is inversely proportional to the temperature.

13.2.4 Space Charge Polarization

Space charge polarization is found in multiphase dielectric substances, where the resistivity changes between the different phases of the dielectric. Due to sudden change in conductivity the electric charges accumulate at the interface when an electric field is applied. The space charge polarization is not an important factor in most dielectrics, and its contribution is not often considered.

13.2.5 Total Polarization

The total polarizability of a dielectric is the sum of the electronic, ionic, and orientation polarizabilities. That is,

$$\alpha = \alpha_e + \alpha_i + \alpha_o$$

$$\alpha = 4\pi\epsilon_0 R^3 + \alpha_i + \frac{\mu^2}{3kT} \quad (13.24)$$

The total polarization \mathbf{P} is,

$$\mathbf{P} = \mathbf{P}_e + \mathbf{P}_i + \mathbf{P}_o$$

$$\mathbf{P} = N \left[4\pi\epsilon_0 R^3 + \alpha_i + \frac{\mu^2}{3kT} \right] \mathbf{E} \quad (13.25)$$

13.3 INTERNAL FIELD IN A DIELECTRIC

Inside a dielectric, a molecule or atom experiences not only the external field applied, but also the fields produced by the dipoles. As coulombic forces have long range, the second contribution is not negligible. The resultant field is called the **local field** or the **internal field**. This local field is responsible for polarizing individual molecules or atoms in solids.

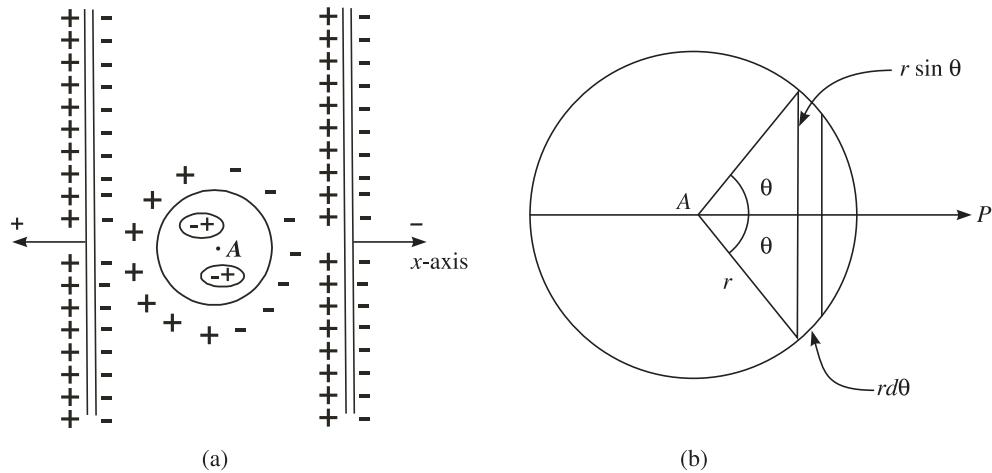


Fig. 13.3 (a) Calculation of internal field in a dielectric showing the charges developed on the plates, on the boundary of the dielectric, and on the surface of imaginary cavity.
(b) Enlarged section of the spherical cavity in a uniformly polarized dielectric.

To calculate the local field, the method suggested by Lorentz is followed. In this method a small imaginary sphere concentric with the reference molecule or atom is considered, as illustrated in Fig. 13.3(a). For a general lattice site, the local field at A is the sum of four contributions.

$$\mathbf{E}_{local} = \mathbf{E}_0 + \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 \quad (13.26)$$

where \mathbf{E}_0 is the field produced by the fixed charges external to the body
 \mathbf{E}_1 is the depolarization field due to the surface charge on the outer surface of the specimen
 \mathbf{E}_2 is the field due to dipoles inside the spherical cavity
 \mathbf{E}_3 is the Lorentz cavity field, i.e. the field due to the polarization charges on the surface of the cavity.

Next, we shall discuss these four contributions:

- (i) Field \mathbf{E}_0 : This field is the one produced by fixed charges external to the body, that is, the field intensity due to the charge density on the outside of the plates. From the constitutive relations in Section 2.8, we have

$$\mathbf{E}_0 = \frac{\mathbf{D}}{\epsilon_0} \quad (13.27)$$

where D is the electric displacement density and ϵ_0 is the permittivity of free space. If \mathbf{E} is the external field applied and \mathbf{P} is the polarization, from Eq. (13.9), we get

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (13.28)$$

Hence,

$$\mathbf{E}_0 = \mathbf{E} + \frac{\mathbf{P}}{\epsilon_0} \quad (13.29)$$

- (ii) Field \mathbf{E}_1 : \mathbf{E}_1 is the field intensity due to the charges induced on the two outer surfaces of the dielectric. From Gauss's law, we get

$$\mathbf{E}_1 = -\frac{\mathbf{P}}{\epsilon_0} \quad (13.30)$$

Since it opposes the field \mathbf{E}_0 , this is often referred to as **depolarization field**.

- (iii) Field \mathbf{E}_2 : Field \mathbf{E}_2 is the one due to the dipoles inside the spherical cavity imagined, which is the only contribution that depends on the crystal structure. For an atom in a cubic site, $\mathbf{E}_2 = \mathbf{0}$ if all the atoms are considered as point dipoles parallel to each other.
- (iv) Field \mathbf{E}_3 : \mathbf{E}_3 is the field intensity due to the polarization charges on the surface of the imagined spherical cavity, which was first calculated by Lorentz. Figure 13.3(b) illustrates an enlarged view of the spherical cavity. Consider the ring of radius $r \sin\theta$.

$$\text{Area of the ring, } dA = 2\pi r \sin\theta r d\theta = 2\pi r^2 \sin\theta d\theta \quad (13.31)$$

The surface density of the charge on the surface of the cavity is equal to the normal component of the polarization, $\mathbf{P} \cos\theta$. Hence, charge on the ring of area dA is:

$$dq = 2\pi r^2 \sin\theta d\theta \mathbf{P} \cos\theta = 2\pi \mathbf{P} r^2 \sin\theta \cos\theta d\theta \quad (13.32)$$

The field at the centre is equal to the force experienced by a unit charge placed at that point. Hence, by Coulomb's law, the field at A is given by

$$d\mathbf{E}_3 = \frac{2\pi \mathbf{P} r^2 \sin \theta \cos \theta d\theta \cos \theta}{4\pi \epsilon_0 r^2} = \frac{\mathbf{P} \cos^2 \theta \sin \theta d\theta}{2\epsilon_0}$$

The factor $\cos \theta$ is introduced to get the component of the field in the direction x. The total field is obtained as

$$\begin{aligned} \mathbf{E}_3 &= \int_0^\pi \frac{\mathbf{P} \cos^2 \theta \sin \theta d\theta}{2\epsilon_0} = -\frac{\mathbf{P}}{2\epsilon_0} \int_0^\pi \cos^2 \theta d(\cos \theta) \\ &= \frac{-\mathbf{P}}{2\epsilon_0} \left[\frac{\cos^3 \theta}{3} \right]_0^\pi = \frac{\mathbf{P}}{3\epsilon_0} \end{aligned} \quad (13.33)$$

Summing up, the total internal field experienced by the dielectric is given by

$$\begin{aligned} \mathbf{E}_{local} &= \mathbf{E} + \frac{\mathbf{P}}{\epsilon_0} - \frac{\mathbf{P}}{\epsilon_0} + \frac{\mathbf{P}}{3\epsilon_0} \\ &= \mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \end{aligned} \quad (13.34)$$

Equation (13.30) is known as **Lorentz relation** for the local field, which is valid for a cubic structure. Substituting the value of \mathbf{P} from Eq. (13.7), we get

$$\mathbf{E}_{local} = \mathbf{E} + \frac{\epsilon_0 \chi \mathbf{E}}{3\epsilon_0} = \left(1 + \frac{\chi}{3}\right) \mathbf{E} \quad (13.34a)$$

The above result has been derived on the basis of simplifying assumptions. Experimentally it has been found that the behaviour of many dielectric materials agrees with the above expression. In general, for any dielectric, the local field that polarizes the dielectric can be expressed as

$$\mathbf{E}_{local} = \mathbf{E} + \frac{\gamma \mathbf{P}}{\epsilon_0} \quad (13.35)$$

where γ is a dimensionless constant whose value depends on the type of the dielectric. This internal field constant is $\frac{1}{3}$ for an isotropic dielectric having cubic symmetry.

13.4 CLAUSIUS–MOSSOTTI RELATION

Let us consider a solid dielectric element, having N atoms per m^3 . \mathbf{E}_l is the local electric field or the internal electric field experienced by the dielectric, when an electric field is applied. The polarization in terms of the local electric field is given as

$$\mathbf{P} = N\alpha \mathbf{E}_l \quad (13.36)$$

where α is the dielectric polarizability. Substituting the value of \mathbf{E}_l from Eq. (13.34), we get

$$\mathbf{P} = N\alpha \left(\mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \right)$$

$$\begin{aligned}\mathbf{P} \left(1 - \frac{N\alpha}{3\epsilon_0}\right) &= N\alpha \mathbf{E} \\ \mathbf{P} &= \frac{N\alpha \mathbf{E}}{\left(1 - \frac{N\alpha}{3\epsilon_0}\right)}\end{aligned}\quad (13.36a)$$

Equating the two expressions for \mathbf{P} , given by Eqs. (13.10) and (13.36a), we get

$$\begin{aligned}\frac{N\alpha \mathbf{E}}{\left(1 - \frac{N\alpha}{3\epsilon_0}\right)} &= \epsilon_0 \mathbf{E}(\epsilon_r - 1) \\ \epsilon_r - 1 &= \frac{N\alpha}{\epsilon_0 \left(1 - \frac{N\alpha}{3\epsilon_0}\right)}\end{aligned}\quad (13.37)$$

Adding 3 on both sides, we get

$$\epsilon_r + 2 = 3 + \frac{N\alpha}{\epsilon_0 \left(1 - \frac{N\alpha}{3\epsilon_0}\right)}\quad (13.38)$$

Dividing Eq. (13.37) by Eq. (13.38), we obtain

$$\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{N\alpha}{3\epsilon_0}\quad (13.39)$$

which relates the relative dielectric constant to the atomic polarizability for crystal structures having cubic symmetry. Equation (13.39) is called the **Clausius—Mossotti relation**.

13.5 FREQUENCY DEPENDENCE OF POLARIZATION

When a dielectric is subjected to an alternating field, the polarization components follow the field reversal at each change in direction of the field. This affects the total polarization, which depends on the ability of dipoles to orient themselves in the field direction during each alternation of the field. The **relaxation time** τ is the average time taken by a dipole to reorient in the field direction, and the reciprocal of the relaxation time is called the **relaxation frequency**.

In the audio frequency region all types of polarization are possible. At low frequencies, dipoles get sufficient time to orient themselves completely along the direction of the field. As the field direction changes, orientation also changes from one direction to another. At frequencies greater than the relaxation frequency of dipoles, the dipoles cannot reverse their direction rapidly, and the total polarization becomes the sum of ionic and electronic polarizations. This phenomenon is more prominent at radio and microwave frequencies.

Due to the inertia of the system, ionic polarization fails to follow field reversal in the infrared region and its contribution stops. So the total polarization is purely electronic polarization. In the optical region, an electron cloud follows the field variations and the material exhibits electronic polarization. In the ultraviolet region, the electron cloud fails to follow the field reversal and the electronic polarization also stops. Figure 13.4 schematically depicts the various contributions to the real part of the polarizability. Permanent dipoles in dipolar materials contribute at low frequencies. The dielectric constant at optical frequencies arises mainly from the electronic polarizability. Therefore, in the optical range $\epsilon_r \approx n^2$, where n is the refractive index, Eq. (13.39) reduces to

$$\frac{n^2 - 1}{n^2 + 2} = \frac{N\alpha_e}{3\epsilon_0} \quad (13.40)$$

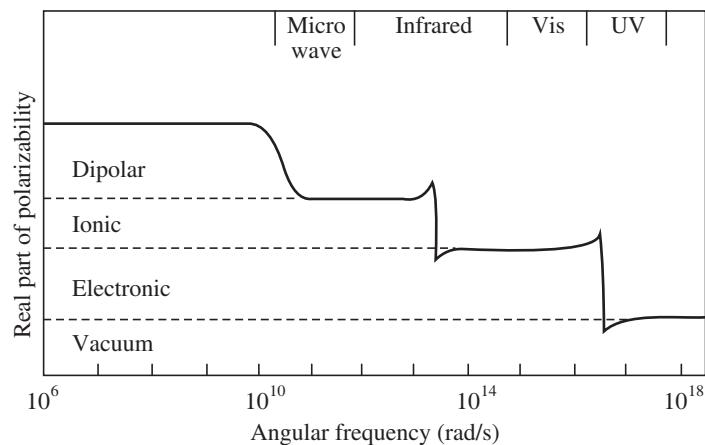


Fig. 13.4 Frequency dependence of polarization.

13.6 DIELECTRIC LOSS

When a dielectric is placed in an electric field, a certain amount of electrical energy is dissipated, which transforms into heat energy. The phenomenon in which electrical energy is absorbed when a dielectric is subjected to an *ac* electric field is known as **dielectric loss**.

Let us consider a capacitor of capacitance C . A dielectric of dielectric constant ϵ is placed between its plates of cross sectional area A , separated by a distance d . Let a sinusoidal voltage V and angular frequency ω be applied between its plates. The dipoles will change their orientation, which will be opposed by the resistance R of the capacitor.

In a resistor, the current I_R and applied voltage are in phase, whereas in a capacitor the current I_c leads the applied voltage V by 90° as shown in Fig. 13.5(a). The resultant current I through a capacitor consists of two components: active or conduction current I_R and displacement or reactive current I_c . Also, the resultant current I lags behind I_c by an angle δ as shown in Fig. 13.5(b). The angle δ is called the **dielectric loss angle**.

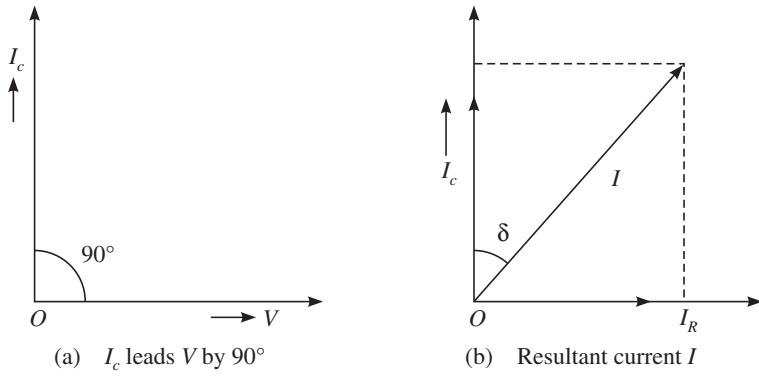


Fig. 13.5 Voltage current phase relationship in a capacitor.

$$\tan \delta = \frac{I_R}{I_c} \quad (13.41)$$

$\tan \delta$ is known as the **loss tangent**, which is also called the **dissipation factor**, as it represents electrical power loss, which is in the form of heat.

For an ideal dielectric, $I_R = 0$ and there is no absorption in electrical energy. The power loss in a real dielectric is given by

$$P_L = VI_R = VI_c \tan \delta \quad (13.42)$$

Displacement current,

$$I_c = \frac{V}{X_c} = \frac{V}{\frac{1}{\omega C}} = \omega CV \quad (13.43)$$

$$P_L = \omega CV^2 \tan \delta$$

$$P_L = 2\pi\nu CV^2 \tan \delta \quad (13.44)$$

where ν is the frequency of the applied voltage. Substituting the value of C for a parallel plate capacitor, we get

$$\begin{aligned} P_L &= 2\pi\nu \frac{\epsilon A}{d} V^2 \tan \delta \\ &= 2\pi\nu \epsilon Ad \frac{V^2}{d^2} \tan \delta \\ P_L &= 2\pi\nu\epsilon\nu E^2 \tan \delta \end{aligned} \quad (13.45)$$

where $\nu = Ad$ represents the volume of the dielectric and $E = \frac{V}{d}$ is the applied electric field.

From Eq. (13.45) it follows that the dielectric loss is directly proportional to the dissipation factor, the dielectric constant, the frequency of the electric field, the electric field, and the volume of the dielectric.

13.7 FERROELECTRIC CRYSTALS

Ferroelectric crystals form an important group of dielectrics which exhibits a state of spontaneous polarization. In the ferroelectric state, centres of positive and negative charges do not coincide and the materials have a net dipole moment even in the absence of an applied electric field. Ferroelectric materials are called **non linear dielectrics** because of their non linear variation of polarization with the external electric field. This non linear variation leads to a closed loop called the **hysteresis loop**, as shown in Fig. 13.6(a).

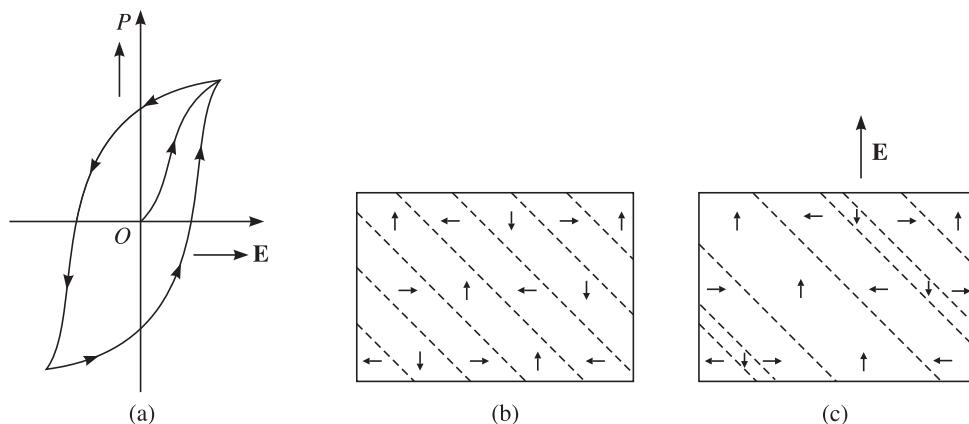


Fig. 13.6 (a) Hysteresis loop of a ferroelectric material, (b) Ferroelectric domains, (c) Ferroelectric domains in the presence of a field.

In ferroelectric crystals there are regions called **ferroelectric domains**, where within a single domain the polarization has a common direction, but in adjacent domains the polarizations are in different directions, as shown in Fig. 13.6(b). The domain boundaries need not be straight lines as shown in the figure. On the application of an electric field, the number and dimensions of the domains polarized in the direction of the electric field may change, as shown in Fig. 13.6(c). The net polarization depends on the difference in the volume of oppositely directed domains. When the volumes of oppositely directed domains are equal, the crystal will appear as unpolarized. The domains exist in the ferroelectric state only in a specific temperature range. When a ferroelectric is heated above a particular temperature called the Curie temperature, the domains disappear and ferroelectric behaviour ceases.

13.8 PIEZOELECTRIC CRYSTALS

When an external stress is applied to certain crystals, electric polarization is produced and a dipole moment is induced. Such crystals are called **piezoelectric crystals** and the phenomenon is called **piezoelectricity**. The piezoelectric effect is reversible. That is, the application of an electric field produces strain in the crystal. Generally, a high value of electric field is required to produce very small strains.

Displacement of ions takes place in certain crystals when stress is applied. This is also regarded as the cause for the occurrence of piezoelectricity. In piezoelectric crystals, the displacement of ions occurs in such a way that the charge distribution loses the original symmetry. This phenomenon is not observed in crystals where the distortion does not change the symmetry of charge distribution. All ferroelectric crystals are piezoelectric, but a crystal may be piezoelectric without being ferroelectric. Quartz is piezoelectric but not ferroelectric, whereas barium titanate is both ferroelectric and piezoelectric.

13.8.1 Applications

Both direct and inverse photoelectric effects are widely used in transducers, strain gauges, and other devices. An alternating potential difference applied to a piezoelectric transducer fixed to a sample causes it to vibrate, mechanically producing sound waves in the sample. A piezoelectric strain gauge is fixed to a sample and the potential difference across it is used to determine the extent of strain in the sample. Piezoelectric crystals are used for the production of ultrasonic waves (Section 17.3). Piezoelectric semiconductors (ZnO, GaS, CdS) are finding applications as amplifiers of ultrasonic waves. The direct piezoelectric effect is used in microphones, piezoelectric compression accelerometers, gas lighters, etc. The inverse piezoelectric effect is used in quartz watches, which keep time accurately. Cooking in microwave oven involves dielectric heating. When wooden sheets are glued, the heat produced in the glue due to dielectric absorption leads to the binding of the wooden sheets.

13.9 APPLICATIONS OF DIELECTRIC MATERIALS

Dielectric materials are mostly used as insulating materials and as medium in capacitors. Some of their applications are listed as follows:

Solid insulating materials

Widely used solid insulators are PVC materials, ceramics, mica, and asbestos. PVC materials have good electrical and mechanical properties. Hence, they are widely used for insulation in wires and cables. PVC films, tapes, and sheets are used as insulators for conductors, cables, and so on.

Ceramics also have excellent dielectric and mechanical properties. In an electric iron, mica or asbestos insulation is provided to isolate current from the outer body of the iron. It is also used as an insulator for switches, plug holders, cathode heaters, and so on.

Asbestos in the form of paper, tape, and board finds extensive use as an insulator. It is also used for the manufacture of panel boards and as insulating tubes in the construction of air cooled transformers, etc.

Liquid insulating material

Liquid dielectrics have cooling ability in addition to insulating properties. Silicon oil, mineral oils, etc. are widely used as filling medium for transformers and circuit breakers. Synthetic oil is used as coolant and insulator in high voltage transformers. Oils having high viscosity and dielectric constant are used for impregnation of papers used in capacitors.

Capacitor material

Mica, glass, etc. are used as dielectric materials in capacitors, since they can withstand very high voltage and temperature. Capacitors with inert gases or air have very low dielectric losses. They are used in radiofrequency circuits. In high voltage applications, capacitors with mineral oils as dielectric are used, since the capacitance values are usually very high. Capacitors with both solid and liquid dielectrics give very high capacitances and are used in electric power distribution system.

Piezoelectric materials

Applications of piezoelectric materials are discussed in Section 13.8.

SOLVED EXAMPLES

Example 13.1 The number of atoms per cubic metre of hydrogen gas is 9.7×10^{26} . The radius of hydrogen atom is 0.52 \AA . Calculate the electronic polarizability.

Solution: From Eq. (13.18), we get

$$\begin{aligned}\alpha_e &= 4\pi\epsilon_0 R^3 \\ &= 4(3.14)(8.85 \times 10^{-12} \text{ F/m})(0.52 \times 10^{-10} \text{ m})^3 \\ &= 1.56 \times 10^{-41} \text{ Fm}^2\end{aligned}$$

Example 13.2 A parallel plate capacitor has an area of 100 cm^2 and a plate separation of 1 cm . If a potential of 100 V is applied to the plates, calculate the capacitance of the capacitor and the charge on the plates.

Solution: The capacitance of the capacitor is given by

$$\begin{aligned}C &= \frac{\epsilon_0 A}{d} \\ &= \frac{(8.854 \times 10^{-12} \text{ F/m})(100 \times 10^{-4} \text{ m}^2)}{10^{-2} \text{ m}} = 8.854 \times 10^{-12} \text{ F}\end{aligned}$$

Since

$$C = \frac{Q}{V}$$

$$\begin{aligned}Q &= CV = (8.854 \times 10^{-12} \text{ F}) 100 \text{ V} \\ &= 8.854 \times 10^{-10} \text{ C}\end{aligned}$$

Example 13.3 A potential of 15 V is applied between the plates of a parallel plate capacitor. A dielectric having relative dielectric constant 5.0 is introduced between the plates. If the separation between the plates is 1.5 mm, calculate the dielectric displacement.

Solution: The Dielectric displacement is given by

$$\begin{aligned} D &= \epsilon_0 \epsilon_r \mathbf{E} = \epsilon_0 \epsilon_r \frac{V}{d} \\ &= \frac{(8.854 \times 10^{-12} \text{ F/m}) \times 5 \times 15 \text{ V}}{1.5 \times 10^{-3} \text{ m}} \\ &= 4.427 \times 10^{-7} \text{ C/m}^2 \end{aligned}$$

Example 13.4 A solid elemental dielectric with number density 3×10^{28} atoms/m³ has an electronic polarizability $\alpha_e = 10^{-40}$ Fm². Calculate the relative dielectric constant of the material.

Solution: For elemental dielectric, the ionic and orientational polarizabilities α_i and α_0 are zero. Hence, $\alpha = \alpha_e$ and from Eq. (13.11a), we obtain

$$\begin{aligned} \epsilon_r &= 1 + \frac{N\alpha_e}{\epsilon_0} \\ &= 1 + \frac{(3 \times 10^{28} / \text{m}^3)(10^{-40} \text{ Fm}^2)}{8.854 \times 10^{-12} \text{ F/m}} \\ &= 1 + 0.339 = 1.339 \end{aligned}$$

Example 13.5 The relative dielectric constant of sulphur is 3.75 and its density is 2050 kg/m³. If the internal field constant $y = \frac{1}{3}$, calculate the electronic polarizability of sulphur. The atomic weight of sulphur is 32.

Solution: Avogadro number $= 6.02 \times 10^{23} / \text{mole}$

$$\text{Number of atoms in one gram} = \frac{6.02 \times 10^{23}}{32}$$

$$\text{Number of atoms in one kilogram} = \frac{6.02 \times 10^{26}}{32}$$

$$\text{Number of atoms/m}^3 = \frac{6.02 \times 10^{26}}{32} \times 2050 = 3.8566 \times 10^{28}$$

Since the internal field constant $y = \frac{1}{3}$, from Eq. (13.35), we get

$$\mathbf{E}_{local} = \mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0}$$

This leads to

$$\begin{aligned}\frac{\epsilon_r - 1}{\epsilon_r + 2} &= \frac{N\alpha_e}{3\epsilon_0} \quad \text{or} \quad \alpha_e = \left(\frac{\epsilon_r - 1}{\epsilon_r + 2} \right) \frac{3\epsilon_0}{N} \\ \alpha_e &= \frac{2.75 \times 3(8.854 \times 10^{-12} \text{ F/m})}{5.75 \times 3.8566 \times 10^{28} / \text{m}^3} \\ &= 3.294 \times 10^{-40} \text{ Fm}^2\end{aligned}$$

Example 13.6 The refractive index of a dielectric material is 1.6. If its number density is 3×10^{28} atoms/m³, calculate its electronic polarizability.

Solution: From Eq. (13.40), we obtain

$$\begin{aligned}\frac{n^2 - 1}{n^2 + 2} &= \frac{N\alpha_e}{3\epsilon_0} \quad \text{or} \quad \alpha_e = \left(\frac{n^2 - 1}{n^2 + 2} \right) \frac{3\epsilon_0}{N} \\ \alpha_e &= \left(\frac{1.6^2 - 1}{1.6^2 + 1} \right) \frac{3(8.854 \times 10^{-12} \text{ F/m})}{3 \times 10^{28} / \text{m}^3} \\ &= 3.03 \times 10^{-40} \text{ Fm}^2\end{aligned}$$

Example 13.7 The refractive index of a dielectric is 1.6 and its relative dielectric constant $\epsilon_r = 4.9$. Estimate the ratio of its electronic polarizability to ionic polarizability.

Solution: From Eq. (13.40), we have

$$\frac{N\alpha_e}{3\epsilon_0} = \frac{n^2 - 1}{n^2 + 2} \quad \text{or} \quad \frac{N}{3\epsilon_0} = \frac{(n^2 - 1)}{(n^2 + 2)} \frac{1}{\alpha_e}$$

From Eq. (13.39), we get

$$\begin{aligned}\frac{\epsilon_r - 1}{\epsilon_r + 2} &= \frac{N(\alpha_e + \alpha_i)}{3\epsilon_0} = \frac{(n^2 - 1)}{(n^2 + 2)} \frac{(\alpha_e + \alpha_i)}{\alpha_e} \\ 1 + \frac{\alpha_i}{\alpha_e} &= \left(\frac{\epsilon_r - 1}{\epsilon_r + 2} \right) \left(\frac{n^2 + 2}{n^2 - 1} \right) \\ 1 + \frac{\alpha_i}{\alpha_e} &= \left(\frac{4.9 - 1}{4.9 + 2} \right) \left(\frac{2.56 + 2}{2.56 - 1} \right) = 1.652 \\ \frac{\alpha_i}{\alpha_e} &= 0.652 \\ \frac{\alpha_e}{\alpha_i} &= 1.53\end{aligned}$$

REVIEW QUESTIONS

- 13.1** Explain polarizability and susceptibility of a dielectric material.
- 13.2** Show that $\epsilon_r - 1 = \frac{N\alpha}{\epsilon_0}$; the symbols have the usual meaning.
- 13.3** Explain electronic polarization and ionic polarization.
- 13.4** What is orientation polarization? State the relation connecting orientation polarizability and temperature.
- 13.5** Explain the concept of internal electric field in dielectrics.
- 13.6** State and explain Clausius–Mossotti relation.
- 13.7** What is relaxation time of polarization? Explain briefly the frequency dependence of polarization in dielectrics.
- 13.8** Represent graphically the variation of the polarizability of a dielectric as a function of the frequency, giving the various contributions and the relevant frequency ranges.
- 13.9** What is dielectric loss? On what factors does it depend?
- 13.10** Explain ferroelectricity. What are ferroelectric domains?
- 13.11** What is piezoelectricity? Explain some of its applications.
- 13.12** Explain polarizability and susceptibility of a dielectric material and derive the relations

$$(i) \quad \epsilon_r - 1 = \frac{N\alpha}{\epsilon_0}$$

$$(ii) \quad \epsilon_r = 1 + \chi .$$

Write a note on dielectric polarization.

- 13.13** What is electronic polarization? Derive an expression for the electronic polarizability of a dielectric material in terms of the radius of the atom.
- 13.14** Explain the origin of the different kinds of polarization in dielectric materials. Discuss the variation of the polarizability of a dielectric as a function of the frequency, giving the various contributions and the relevant frequency ranges.
- 13.15** Explain the concept of internal electric field in dielectrics and derive an expression for the internal field of a dielectric having cubic symmetry.
- 13.16** Outline the different contributions for the internal electric field in a dielectric. State the relation for the local field in a dielectric. Derive the Clausius–Mossotti relation for a dielectric having cubic symmetry.

PROBLEMS

1. A parallel plate capacitor has an area of 6 cm^2 and a plate separation of 0.5 cm . The charge on the plate is $2 \times 10^{-10}\text{C}$. When a material of relative dielectric constant 3.0 is introduced between the plates, what is the voltage across the capacitor plates?
2. The relative dielectric constant of He gas is 1.0000684. If the gas contains $2.75 \times 10^{25} \text{ atoms/m}^3$, calculate the electronic polarizability of He atoms.
3. Consider a parallel plate capacitor with amorphous selenium of relative dielectric constant $\epsilon_r = 6.0$ between the plates. If the concentration of selenium is $3.67 \times 10^{28} \text{ atoms/m}^3$, estimate the polarizability of the selenium atom.
4. The radius of the helium atom is 0.55 \AA . Calculate its polarizability. If the number density of helium atoms is $2.7 \times 10^{25}/\text{m}^3$, what is its relative dielectric constant?

C H A P T E R

14

Magnetic Properties of Materials

Researching and studying the properties of magnetic substances have increased in the recent years although the subject is more than 2000 years old. Magnetism plays a very important role in many modern technologies. The revolution in information technology is partly due to the developments in magnetic storage of information. In this chapter, we shall discuss some of the basic ideas regarding the origin of atomic magnetism, the different types of magnetic materials and their properties.

14.1 PERMEABILITY AND SUSCEPTIBILITY

We can imagine that a magnetic field surrounds every magnet. A magnetic field can be described by either the **magnetic induction \mathbf{B}** (or magnetic flux density) or the **magnetic field intensity \mathbf{H}** . The magnetic induction is defined by the number of magnetic lines of force passing perpendicularly through unit area. Its unit in S.I is Tesla (Wbm^{-2}). The magnetic field intensity at any point in a magnetic field is the force experienced by a unit north pole placed at that point. Its unit is Am^{-1} . In vacuum \mathbf{B} and \mathbf{H} are related by the equation

$$\mathbf{B} = \mu_0 \mathbf{H} \quad (14.1)$$

where $\mu_0 = 4\pi \times 10^{-7} \text{ NA}^{-2}$ (TmA^{-1}) and is called the **permeability of free space** (vacuum).

When a substance is placed in a magnetic field, it gets magnetized. Consequently, magnetization \mathbf{M} which is defined as the magnetic moment per unit volume, is produced in it. Its unit is Am^{-1} . The magnetic induction inside the substance is given by

$$\mathbf{B} = \mu_0 \mathbf{H} + \mu_0 \mathbf{M} = \mu_0 (\mathbf{H} + \mathbf{M}) \quad (14.2)$$

Here the first term is due to the external magnetic field and the second term is due to the magnetization of the sample. For an isotropic medium, magnetization \mathbf{M} is proportional to \mathbf{H}

$$\mathbf{M} = \chi \mathbf{H} \quad (14.3)$$

where the scalar quantity χ is called the **susceptibility** of the medium. Substituting the value of \mathbf{M} in Eq. (14.2), we get

$$\mathbf{B} = \mu_0 \mathbf{H}(1 + \chi) = \mu \mathbf{H} \quad (14.4)$$

where

$$\mu = \mu_0(1 + \chi) \quad (14.5)$$

is known as the **permeability** of the medium. Often, it is more convenient to use the relative permeability μ_r defined by

$$\mu_r = \frac{\mu}{\mu_0} = 1 + \chi \quad (14.6)$$

14.2 ORIGIN OF MAGNETIC MOMENT

14.2.1 Orbital Magnetic Moment

As a rough model, atoms consist of electrons circulating around the central positive nucleus. The moving electrons give rise to an angular momentum, called **orbital angular momentum** \mathbf{L} , to the atom. Considering the electrons as current loops, we can derive an expression for the magnetic moment μ_L associated with the angular momentum. The current I due to the orbital motion of an electron of charge e is given by

$$I = \frac{e}{T} \quad (14.7)$$

where T is the period of revolution. If r is the radius of the orbit and v is its velocity

$$T = \frac{2\pi r}{v} \quad (14.8)$$

According to Ampere's law, the magnetic moment due to this current is given by

$$\begin{aligned} \mu_L &= I \times \text{Area of the orbit} = \frac{e}{T} \times \pi r^2 \\ &= \frac{2\pi r}{T} \frac{er}{2} = \frac{evr}{2} = \frac{e}{2m} mvr \\ &= \frac{e}{2m} L \end{aligned} \quad (14.9)$$

where $L = mvr$ is the magnitude of the orbital angular momentum. The expression for μ comes in a more elegant way in quantum mechanics. The magnetic moment μ_L associated with an electron having charge e is given by

$$\mu_L = -\frac{e}{2m} \mathbf{L} \quad (14.10)$$

where m is the mass of the electron. The negative sign indicates that the vectors μ_L and \mathbf{L} are oppositely directed. The magnitude of \mathbf{L} according to quantum mechanics is:

$$|\mathbf{L}| = \sqrt{l(l+1)} \hbar \quad (14.11)$$

where the quantum number l , called the orbital quantum number, can take the values

$$l = 0, 1, 2, \dots, (n-1) \quad (14.12)$$

where n is the total quantum number of the electron in the atom. Substituting this value of \mathbf{L} in Eq. (14.10), we get

$$|\mu_L| = \frac{e}{2m} \sqrt{l(l+1)} \hbar = \mu_B \sqrt{l(l+1)} \quad (14.13)$$

where

$$\mu_B = \frac{e\hbar}{2m} \quad (14.14)$$

is called **Bohr magneton**. Its value is

$$\mu_B = 9.27 \times 10^{-24} \text{ JT}^{-1} (\text{Am}^2) \quad (14.15)$$

14.2.2 Spin Magnetic Moment

To explain the results of optical spectra and the results of Stern–Gerlach experiment, Uhlenbeck and Goudsmit proposed that the electron in an atom has an additional intrinsic angular momentum, called **spin angular momentum \mathbf{S}** , specified by the spin quantum number s . They also suggested that \mathbf{S} is similar to the vector \mathbf{L} and has the magnitude

$$|\mathbf{S}| = \sqrt{s(s+1)} \hbar \quad (14.16)$$

Here s is the spin quantum number, which is $\frac{1}{2}$ for an electron. The magnetic moment associated with the spin angular momentum is:

$$\mu_s = -\frac{e}{m} \mathbf{S} \quad (14.17)$$

This result was theoretically confirmed by Dirac in his relativistic quantum mechanics. Substituting the value of \mathbf{S} , we have

$$|\mu_s| = \frac{e}{m} \sqrt{s(s+1)} \hbar = 2\mu_B \sqrt{s(s+1)} \quad (14.18)$$

14.2.3 Total Magnetic Moment

An electron in an atom has both orbital and spin magnetic moments. These two magnetic moments together give the total magnetic moment as

$$\mu = \mu_L + \mu_s \quad (14.19)$$

$$= \mu_B \left[\sqrt{l(l+1)} + 2\sqrt{s(s+1)} \right] \quad (14.20)$$

In many electron systems, different ways of combining the orbital and spin contributions are possible depending on the interaction between the two.

14.3 CLASSIFICATION OF MAGNETIC MATERIALS

Most of the atoms have an incompletely filled shell structure, and therefore have a magnetic moment. Atoms with filled subshells have a zero angular momentum, and therefore no magnetic moment. In ionic materials, the positive ions lose as many electrons to the negative ions so that the resulting ions have only filled subshells. In covalent materials, each covalent bond is formed by a pair of electrons with opposite spins with a net orbital angular momentum of zero. Although a hydrogen atom has a finite magnetic moment, a hydrogen molecule does not. Similarly covalent solids also have zero magnetic moment. In other words, in a large number of solids, the resultant magnetic moment is zero.

Materials may be classified into three types depending on the sign and magnitude of the susceptibility. Materials for which the susceptibility χ is negative are known as **diamagnetic**. A diamagnetic solid has a tendency to repel the magnetic lines of force. Paramagnetism and ferromagnetism are the result of intrinsic magnetic moment. Some atoms and ions possess permanent magnetic moment. In the absence of an external magnetic field, the substance exhibit no net magnetic moment. In the presence of a field the moments tend to align along the field producing a net magnetization. If the susceptibility χ is small and positive, the substance is **paramagnetic**. The paramagnetic susceptibility decreases with increase of temperature. For the third category, called **ferromagnetic**, χ is very large and it becomes spontaneously magnetized below a certain temperature. The high value of χ is partly due to the ordering of adjacent magnetic dipoles in the same direction. Hence, from Eq. (14.4), $B_{dia} < \mu_0 H$. For paramagnetic and ferromagnetic substances $B_{para} > \mu_0 H$ and $B_{ferro} \gg \mu_0 H$. This aspect of the magnetic induction B is illustrated in Fig. 14.1.

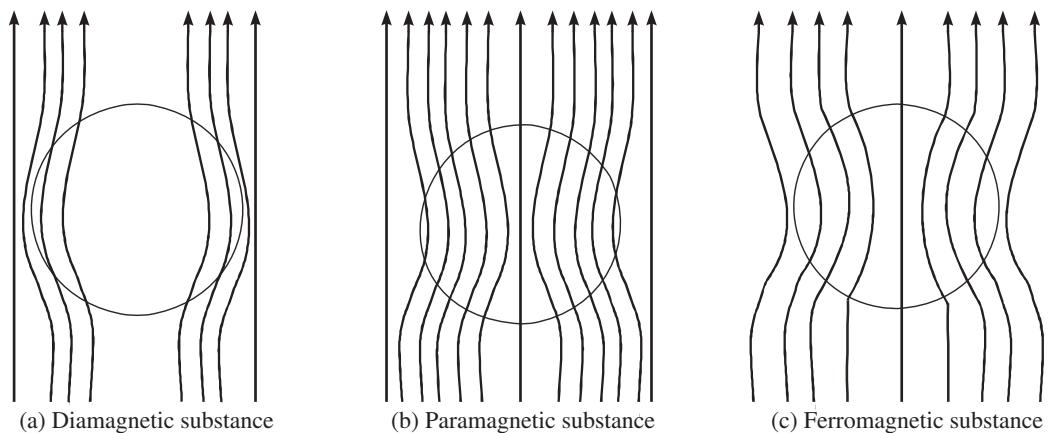


Fig. 14.1 Magnetic induction in a magnetic field.

Based on the ordering of adjacent dipoles, there are two other principles: antiferromagnetism and ferrimagnetism. In antiferromagnetism, the adjacent dipoles are antiparallel, whereas in ferrimagnetism the adjacent dipoles are anti-parallel, but of unequal magnitude. We shall discuss these phenomena in the subsequent sections.

14.4 DIAMAGNETISM

Diamagnetic substances have atoms or ions or molecules with filled subshells, and therefore, possess zero magnetic moment. However, the presence of external magnetic field distorts the orbital motion of the electrons in the solid. Consequently, in the presence of the field, the electrons in an atom generate a magnetic field which opposes the external field applied. This is referred to as **diamagnetism** which occurs in all types of atom. It produces a finite magnetic susceptibility in solids. Some common examples of diamagnetic substances include inert gases, bismuth, copper, zinc, gold, and so on. As the induced field is opposite to the field applied, the susceptibility is negative. The susceptibility does not vary with temperature. Table 14.1 gives susceptibilities for some representative diamagnetic substances.

Table 14.1 Magnetic susceptibility of certain diamagnetics

Material	Susceptibility
Copper	-9.8×10^{-6}
Gold	-3.6×10^{-5}
Silver	-2.6×10^{-5}
Lead	-1.7×10^{-5}
Mercury	-3.2×10^{-5}

A theoretical expression for diamagnetic susceptibility was proposed by Langevin:

$$\chi = -\frac{\mu_0 e^2 N Z}{6m} \langle r^2 \rangle \quad (14.21)$$

where N is the number of atoms per unit volume, Z is the atomic number of the atom, $\langle r^2 \rangle$ is the mean square distance of the electron from the nucleus, e and m are respectively the charge and mass of the electron. The agreement of the calculated values with experimental results is generally satisfactory.

14.5 PARAMAGNETISM

Paramagnetism is exhibited by a substance with unfilled electron shells. In the absence of a magnetic field such atoms and ions exhibit no magnetic moment. Following are some of the systems with unfilled electron shells:

- (i) Atoms and molecules having an odd number of electrons, since the total spin of such systems is nonzero
Examples: Alkali atoms, nitric oxide, organic free radicals

- (ii) Atoms and ions with a partly filled inner shell
Examples: Transition elements (Mn^{2+} , Gd^{3+} , U^{4+}), rare earths, actinide group elements
- (iii) Certain molecules and compounds with even number of electrons
Examples: Molecular oxygen, boron, organic biradicals.

In simple metals like sodium or aluminum, the valence electrons are no longer attached to any particular atom. As a result, although the metal ions do not contribute to the magnetic moment, the moving electrons produce a non zero magnetic moment. This free electron paramagnetism is a weak phenomenon and is often referred to as **Pauli paramagnetism** or paramagnetism of conduction electrons. In an applied magnetic field, these magnetic moments align in the direction of the field, and therefore the Pauli paramagnetism gives a positive contribution to χ . For many metals such as copper, silver and gold, the diamagnetic contribution is larger in magnitude than the paramagnetic contribution. Therefore, these materials display diamagnetic properties. The Pauli paramagnetism is also constant with temperature. Pauli, using the Fermi-Dirac distribution for the free electrons, derived the expression for susceptibility

$$\chi = \frac{\mathbf{M}}{\mathbf{H}} = \frac{\mu_0 N \mu^2}{k T_F} \quad (14.22)$$

where T_F is the Fermi temperature defined by Eq. (10.43), k is the Boltzmann constant, and N is the number of atoms per unit volume. It is obvious from Eq. (14.22) that the susceptibility χ is independent of temperature. The observed values are in broad agreement with the values given by Eq. (14.22).

There are other paramagnetic materials which show strong magnetic behaviour. In transition metals, rare earths and actinide group, the electronic structure is very different from that of the simple metals. In the iron group transition elements, the 4s state is filled before the 3d ones are filled. The inner electrons form an incomplete shell and give rise to a permanent magnetic moment. In other words, the transition elements are capable of producing a significant susceptibility. In the other transition metals groups, it is the 4d and 5d sub shells that are incomplete. In the rare earths and actinide group, it is the 4f and 5f sub shells that are not filled. These elements give rise to what is called as **Curie paramagnetism**.

Based on quantum mechanical considerations, it can be shown that this susceptibility is inversely proportional to the temperature. That is

$$\chi = \frac{\mu_0 N \mu^2}{3kT} = \frac{C}{T} \quad (14.23)$$

$$C = \frac{\mu_0 N \mu^2}{3k} = \frac{\mu_0 N p^2 \mu_B^2}{3k} \quad (14.24)$$

where C is a constant called **Curie constant** and p is the effective number of Bohr magnetons. This relationship is known as **Curie's law**. In the diamagnetism or Pauli paramagnetism, a weak magnetic response is experienced. In Curie paramagnetism, the susceptibility is much larger; however, an external magnetic field is still required to align the magnetic dipoles. The variation of susceptibility of diamagnetic and paramagnetic materials with temperature is shown in Fig. 14.2.

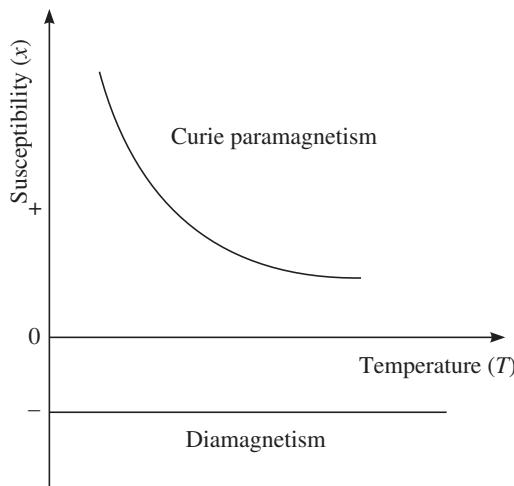


Fig. 14.2 Variation of magnetic susceptibility of diamagnetic and paramagnetic materials with temperature.

14.6 FERROMAGNETISM

Paramagnetism and ferromagnetism are the results of intrinsic magnetic moment. In the presence of a field, the moments tend to align along the field, producing magnetization. If the magnetization is large and exists even in the absence of the field, the substance is ferromagnetic. Again, consider the iron group transition metals. In this group, the incomplete 3d subshell electrons interact strongly with other ions, giving rise to two effects on the magnetic moment of the ions.

- (i) The interaction affects the orbital angular momentum of the 3d electrons in such a way that the average orbital angular momentum on neighbouring ions gets cancelled. In other words, the magnetic moment due to orbital angular momentum becomes zero (quenched).
- (ii) The spins of the electrons interact in such a way that there is correlation between the spins of the 3d electrons on neighbouring ions. This is called Heisenberg's **exchange interaction**. It is a very short range interaction.

14.6.1 Exchange Interaction

To learn more about the exchange interaction, let us consider two adjacent ions each of which has a single 3d electron. If the interaction between the two electrons is sufficiently strong, which is the case in certain materials, the spins of these two electrons will be aligned in the same direction as in Fig. 14.3(a). Since the interaction is of a shorter range, only spins of electrons in the adjacent ions are altered. Thus, each ion alters the orientation of the dipole moment of its neighbouring ions. This results in the alignment of all the atomic dipoles in the crystal as in Fig. 14.3(b). Such a material is called a ferromagnetic material. A ferromagnetic substance can have large magnetization even in the absence of an external magnetic field.

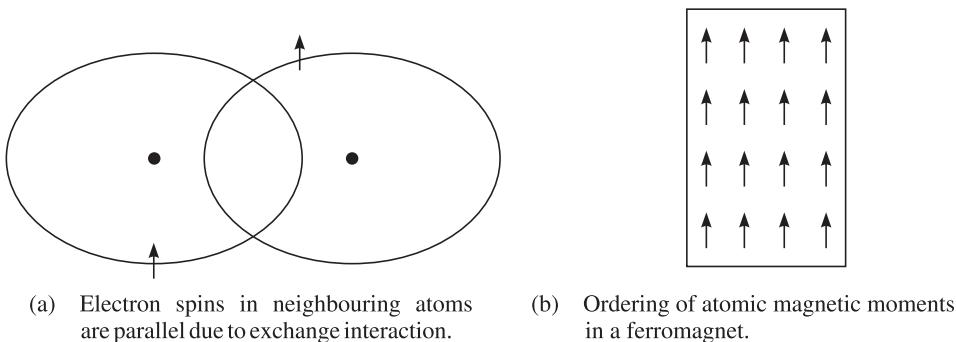


Fig. 14.3 Exchange interaction in ferromagnetic materials.

14.6.2 Curie Temperature

The magnetic susceptibility of ferromagnetic materials is very large. The best known examples of ferromagnets are the transition metals Fe, Co, Ni, their alloys, and the rare earth metals Gd and Dy. Ferromagnetism appears only below a certain temperature called the **ferromagnetic transition temperature** or simply the **Curie temperature** (T_c). This temperature depends on the substance and is well above the room temperature. Table 14.2 gives T_c values along with the saturation magnetization for certain ferromagnetic substances. Above the Curie temperature, the magnetic moments are oriented randomly, resulting in zero net magnetization. In this region the substance is paramagnetic and its susceptibility is given by **Curie–Weiss law**:

$$\chi = \frac{C}{T - T_c} \quad (14.25)$$

The constant C is the Curie constant. Equation (14.25) is of the same form as Eq. (14.23) except that the origin of temperature is shifted from absolute zero to Curie temperature T_c .

Table 14.2 Curie temperature (T_c) and saturation magnetization (M_s) at 300 K for ferromagnetic substances

Substance	T_c (K)	M_s ($\times 10^5$ JT $^{-1}$ m $^{-3}$)	$\mu_0 M_s$ (T)
Iron	1043	17.1	2.15
Cobalt	1388	14.0	1.76
Nickel	627	4.85	0.61
Gadolinium*	292	20.6	2.60
CrO_2	386	5.18	0.65

*For Gd, the M_s value is at 0 K.

In the temperature range $T < T_c$, the spontaneous magnetization, often referred to as the **saturation magnetization**, increases as the temperature is lowered, reaching its maximum at $T = 0$ K. Figure 14.4 shows the ratio of saturation magnetization at temperature T to that at

0 K versus $\frac{T}{T_c}$. The curve is almost the same for Fe, Co, and Ni. As the temperature is reduced, more and more dipoles line up in the magnetization direction.

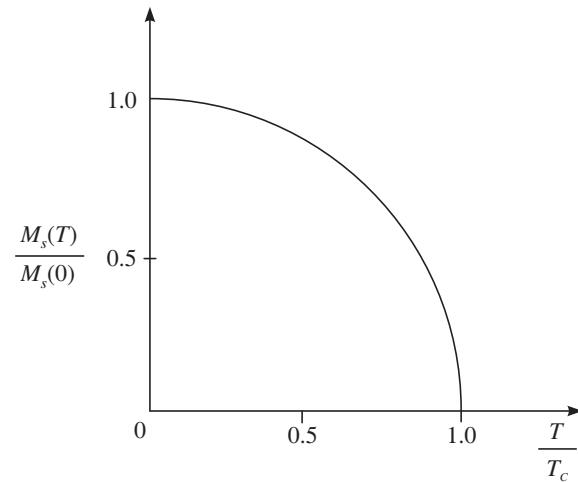


Fig. 14.4 Ratio of saturation magnetization at temperature T to that at 0 K.

14.7 WEISS THEORY OF FERROMAGNETISM

Ferromagnetic materials are not always magnetized even when the temperature is lower than the Curie temperature. To explain this, Weiss postulated that the substance is divided into many small domains. Each domain is magnetized, but the direction of magnetization in the various domains are different. Since the domains are randomly oriented, the net magnetization of the ferromagnetic material is zero. This is shown schematically in Fig. 14.5. The volume of domains

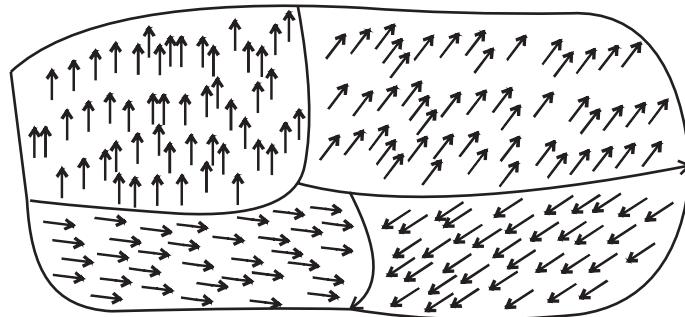


Fig. 14.5 Schematic representation of the domains in a ferromagnet.

is typically of the order of $10^{-8} - 10^{-12} \text{ m}^3$. Between adjacent domains there is a narrow transition region in which each dipole is slightly misaligned with the neighbouring dipole (Fig. 14.6). This

region is called a **Bloch wall**. The width of this region is typically a few hundred atomic spacings.

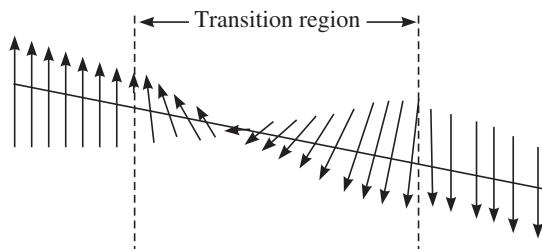


Fig. 14.6 Illustration showing the orientation change from one domain to the next.

14.8 THE HYSTERESIS LOOP

Although a ferromagnet is defined to exhibit magnetization in the absence of an external magnetic field, a sample of a finite size usually exhibits zero magnetization below its Curie temperature. In order to magnetize it, we have to apply an external field. When the field is applied, the domain walls move and more of the dipoles align with the field. Consequently the material gets magnetized. Figure 14.7 illustrates the progress of the magnetization process as the external field \mathbf{H} increases. Starting at the origin, the magnetization \mathbf{M} increases as the field is increased, and finally saturates at point A. The saturation magnetization M_s corresponds to the situation wherein all the magnetic dipoles are aligned with the external field. If the field is now reduced, the new curve does not retrace the original curve OA; it follows the path AB as shown in Fig. 14.7. In other words, the domain walls do not return to their original positions, and therefore the material remains magnetized.

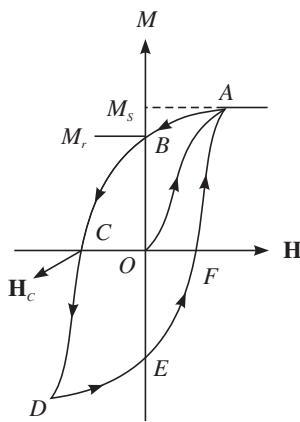


Fig. 14.7 A typical hysteresis curve.

The residual magnetization when the external field has been reduced to zero is called the **remanent magnetization** M_r . To destroy the magnetization to zero, a magnetic field in the

opposite direction is needed. This negative field is called the **coercive field** (H_c). If the field is reduced further and then in the original direction, the complete curve is obtained. If the field alternates periodically, the magnetization traces the solid curve as shown in Fig. 14.7, known as the **hysteresis loop**. As magnetization is related to magnetic induction \mathbf{B} by Eq. 14.4, instead of \mathbf{M} versus \mathbf{H} , we can also plot the variation of \mathbf{B} against \mathbf{H} .

Hysteresis implies the existence of energy losses in the system. The area enclosed by the loop represents the loss of energy in taking the ferromagnetic sample through a complete cycle of magnetization. This is called the **hysteresis loss**. Based on hysteresis loss, magnetic materials are classified into hard and soft magnetic materials.

14.9 SOFT AND HARD MAGNETIC MATERIALS

The hysteresis loop is very important, as the product MH has a lot of practical applications. The materials that give rise to a larger area of the hysteresis loop are called **hard magnetic materials** and those with smaller areas are referred to as **soft magnetic materials**. The hysteresis loops of hard and soft magnetic materials are shown in Fig. 14.8.

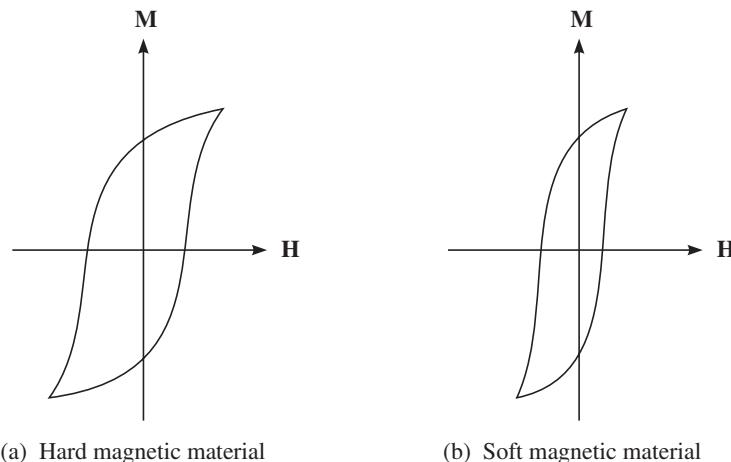


Fig. 14.8 Typical hysteresis curves.

For hard materials a large remanent magnetization and large coercive force is required. They are usually made from fine grained samples (typically 10 nm to 100 nm) so that each particle has a single domain. These materials have small values of permeability and susceptibility. Soft materials have low coercive force and high permeability. They are made from large crystal grains to minimize the interaction between the domain walls. Examples of some hard and soft magnetic materials are listed below:

Hard magnetic materials:

- (i) Aluminium - nickel - cobalt - copper - iron alloys (Alnico V)
- (ii) Copper - nickel - iron alloy (Cunife)
- (iii) Copper - nickel - cobalt alloy (Cunico)

(iv) Sm - Co alloy

(v) $\text{Nd}_2\text{Fe}_{14}\text{B}$

The last two have very high value for the MH energy product

Soft magnetic materials:

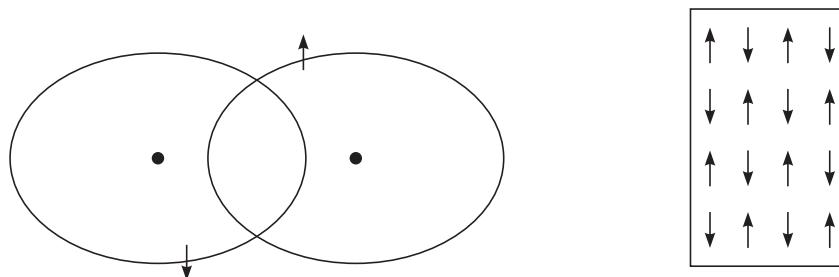
- (i) Ferrous - nickel alloy (80 Ni, 20 Fe and 45 Ni, 55 Fe)
- (ii) Iron - silicon alloy (97 Fe, 3 Si)
- (iii) Aluminium - silicon - iron alloy (5 Al, 10 Si, 85 Fe)
- (iv) Ferrites (Ni - Mn; Mn - Zn)
- (v) Garnet ($\text{Y}_3\text{Fe}_5\text{O}_{12}$; Y-Al-Gd ; Y-Al)

Hard magnetic materials are generally used for the construction of permanent magnets whereas soft magnetic materials are used for transformer cores.

Permanent magnets are used in dynamos, motors, generators, galvanometers, ammeters, voltmeters, tape recorders, TV tubes, laboratory transformers, etc. Soft material magnets are extensively used in chokes, input transformers, output transformers, magnetic modulators, magnetic shields, recording head, magnetic amplifier, etc.

14.10 ANTIFERROMAGNETISM

The concept of exchange interaction occurs only in a limited number of materials. In certain others, the exchange interaction causes the electrons in adjacent ions to align antiparallel, as shown in Fig. 14.9(a). When this is extended to the entire crystal the magnetic dipoles are arranged as shown in Fig. 14.9(b). The neighbouring dipoles cancel each other, leading to a net magnetization of almost zero. Such materials are called **antiferromagnets**. Antiferromagnetism is exhibited by many compounds involving transition metals such as MnF_2 , MnO , FeO , NiO , CoO , Cr_2O_3 . Antiferromagnetism disappears at a certain temperature, called **Neel temperature** (T_N), above which the substance is paramagnetic. A large number of compounds that exhibit



(a) Electron spins in neighbouring atoms are antiparallel due to exchange interaction

(b) Ordering of atomic magnetic moments in an antiferromagnetic material.

Fig. 14.9 Antiparallel alignment in an antiferromagnetic material.

Curie paramagnetism at room temperature are found to become antiferromagnetic at lower temperatures. The variation of susceptibility with temperature is shown in Fig. 14.10(a).

In Fig. 14.10(b), the variation of $\frac{1}{\chi}$ with temperature T is illustrated for paramagnetic, ferromagnetic, and antiferromagnetic substances.

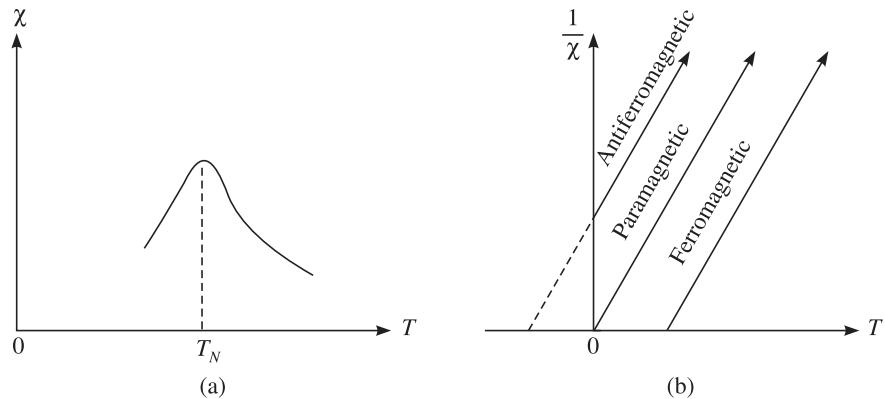


Fig. 14.10 (a) Variation of susceptibility of an antiferromagnetic substance with temperature under a magnetic field, (b) Variation of $\frac{1}{\chi}$ with temperature for paramagnetic, ferromagnetic and antiferromagnetic substances.

Antiferromagnetic crystals can be considered to be composed of two interpenetrating sublattices, say A and B . The atoms of one of the sublattices have their spin oriented in one direction while the spins of the other are in the opposite direction. This model is called the **two sublattice model**. The arrangement is illustrated in Fig. 14.11.

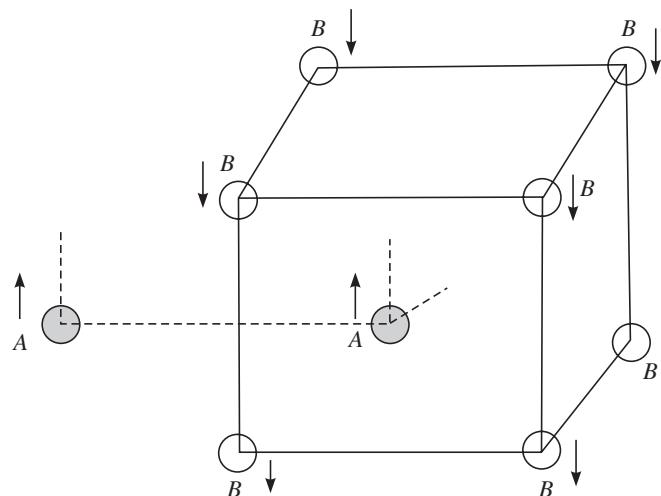


Fig. 14.11 The two sublattices model of an antiferromagnet. The spin of B is downward. A represents the other sublattice whose spin is upward.

14.11 FERRIMAGNETISM

There is yet another way of ordering the atomic dipoles. In a number of materials, the magnetic dipoles are arranged in an antiparallel formation, but the magnitudes of the dipoles pointing in one direction are larger than those aligned in the opposite direction, as illustrated in Fig. 14.12(a). Since the adjacent magnetic moments are antiparallel but of unequal magnitudes, it results in a finite magnetization.

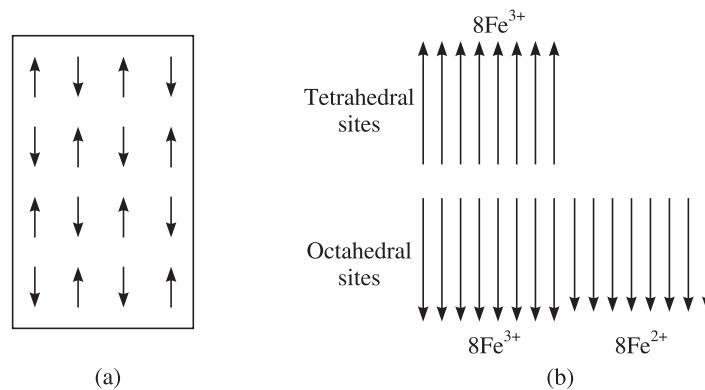


Fig. 14.12 (a) Ordering of atomic magnetic moments in ferrimagnetic material, (b) Spin arrangements in $\text{Fe}^{2+} + \text{Fe}^{3+} \text{O}_4$.

Such materials are called **ferrimagnetic**, a word coined by Neel. The non-vanishing net magnetization arises due to several reasons. However, the two major magnetization that arises are:

- The two sublattices are occupied by different types and numbers of magnetic ions.
- The two sublattices correspond to two different crystallographic sites.

Above a critical temperature called the **ferrimagnetic Neel temperature** T_{FN} , a ferrimagnet becomes a paramagnet.

14.11.1 Ferrites

Ferrimagnetic substances, often referred to as **ferrites**, are ionic oxide crystals having chemical composition of the form $X\text{Fe}_2\text{O}_4$, where X is a divalent metal. The most familiar example of this type is magnetite (lodestone), whose chemical formula is Fe_3O_4 . The unit cell is a close packed cubic structure containing 56 atoms, 32 oxygen atoms, and 24 Fe. Out of the 24 Fe, sixteen are Fe^{3+} and eight Fe^{2+} (in general divalent metal). When the ferrous ion is replaced by the divalent metal ion such as Mn, Co, or Ni, the resulting ferrite has the form $\text{Me}^{2+}\text{Fe}^{3+}\text{O}_4$ ($\text{Me} = \text{Mn, Co, Ni}$). Such ferrites are called **mixed ferrites**. The spin states of Fe^{3+} and Fe^{2+} ions of the $\text{Fe}^{2+}\text{Fe}^{3+}\text{O}_4$ crystal are illustrated in Fig. 14.12(b). Eight of the Fe^{3+} are spin up ones and the remaining eight spin down ones, leading to a net spin of zero. Hence, the contribution of Fe^{3+} to magnetic moment is zero. However, all the eight Fe^{2+} contribute a net magnetic moment. As the structure of ferrites are different from those of ferromagnets, the

temperature dependence of magnetic properties are different from those of ferromagnets. Ferrites are the most useful of all magnetic materials because they are also good electrical insulators.

As ferrites have very high d.c. resistivity, 10^{10} times the resistivity of ferromagnetic iron, eddy current losses will be very low. Hence, they find extensive use in high frequency applications. They are widely used as core materials to increase flux linkage and inductance in high frequency inductors and transformers. The hysteresis loop of spinel ferrites are rectangular in shape and are used as computer memory core elements.

14.12 MAGNETIC MATERIALS FOR INFORMATION STORAGE

The important application of magnetic materials is in information storage devices such as audio and video tapes, floppy discs, and hard discs for computers. Whether the data stored is digital (computer discs) or analog (audio and videotapes), the basic principle is the same. In each case the magnetic medium is covered by many small magnetic particles, each of which behaves as a single domain. A magnet called the **write head** is used to record information to this medium by orienting the magnetic particles along particular directions. Reading the information from the tape or disc is simply the reverse of this process in which the magnetic field produced by the magnetic particles induces a magnetization in the read head.

The material for the magnetic recording medium must have a Curie temperature well above the temperature to which the material will normally be exposed. In addition, the coercivity must be low enough so that the orientation of the particles can be altered by the write head, but high enough so that the orientation is not accidentally affected by external magnetic fields. Magnetic hard discs are used in computers because they are considerably cheaper.

A more recent approach to magnetic data storage is to use magneto-optic discs. In this approach, the write process is done by using a laser beam to heat a magnetic particle above the Curie temperature when a magnetic field is on. To read, another polarized laser beam is directed on to the surface. From the polarization of the reflected beam, the information is obtained.

SOLVED EXAMPLES

Example 14.1 The density and atomic weight of a substance are 7.9 g/cm^3 and 56, respectively. Calculate its spontaneous magnetization if its magnetic moment is $2.2 \mu_B$. The value of Bohr magneton = $9.27 \times 10^{-24} \text{ JT}^{-1}$ (or Am^2).

$$\begin{aligned}\textbf{Solution: } \text{The number of atoms in 1 g of the substance } & \frac{N}{A} = \frac{6.02 \times 10^{23}}{56} \\ & = 1.075 \times 10^{22} / \text{g}\end{aligned}$$

$$\text{The number of atoms in unit volume } n = 7.9 \times 1.075 \times 10^{22}$$

$$= 8.49 \times 10^{22} / \text{cm}^3 = 8.49 \times 10^{28} / \text{m}^3$$

Magnetization $\mathbf{M} = n \times$ Magnetic moment

$$\begin{aligned}&= (8.49 \times 10^{28} \text{ m}^{-3}) 2.2(9.27 \times 10^{-24} \text{ Am}^2) \\&= 1.73 \times 10^6 \text{ Am}^{-1}\end{aligned}$$

Example 14.2 Find the relative permeability of a ferromagnetic material if a field strength of 200 Am^{-1} produces a magnetization of 3100 Am^{-1} .

Solution: Relative permeability $\mu_r = 1 + \chi$

$$\begin{aligned}\text{Susceptibility } \chi &= \frac{\mathbf{M}}{\mathbf{H}} = \frac{3100 \text{ Am}^{-1}}{200 \text{ Am}^{-1}} \\&= 15.5 \\&\mu_r = 16.5\end{aligned}$$

Example 14.3 A field strength of 300 Am^{-1} produces a magnetization of 4400 Am^{-1} in a ferromagnetic material. What is the relative permeability of the material?

$$\text{Susceptibility } \chi = \frac{\mathbf{M}}{\mathbf{H}} = \frac{4400 \text{ Am}^{-1}}{300 \text{ Am}^{-1}} = 14.67$$

$$\begin{aligned}\text{Relative permeability } \mu_r &= 1 + \chi \\&= 1 + 14.67 = 15.67\end{aligned}$$

Example 14.4 The magnetic field in a diamagnetic material is 10000 Am^{-1} . Calculate the magnetization and flux density of the material if its susceptibility is -0.4×10^{-5} .

$$\begin{aligned}\text{Magnetization } \mathbf{M} &= \chi \mathbf{H} = (-0.4 \times 10^{-5})(10000 \text{ Am}^{-1}) \\&= -0.04 \text{ Am}^{-1}\end{aligned}$$

$$\begin{aligned}\text{Flux density } \mathbf{B} &= \mu_0(\mathbf{H} + \mathbf{M}) \\&= (4\pi \times 10^{-7} \text{ TmA}^{-1})(10000 - 0.04) \text{ Am}^{-1} \\&= 0.0126 \text{ T}\end{aligned}$$

Example 14.5 The magnetic field in a diamagnetic material is $1.2 \times 10^5 \text{ Am}^{-1}$. Calculate the magnetization and flux density in it, if its magnetic susceptibility is -4.2×10^{-6} . Also, compute the relative permeability of the material.

Solution: Magnetization, $\mathbf{M} = \chi \mathbf{H}$

$$\begin{aligned}&= (-4.2 \times 10^{-6})(1.2 \times 10^5 \text{ Am}^{-1}) \\&= -0.504 \text{ Am}^{-1}\end{aligned}$$

$$\begin{aligned}\text{Flux density, } \mathbf{B} &= \mu_0(\mathbf{H} + \mathbf{M}) \\&= (4\pi \times 10^{-7} \text{ TmA}^{-1})(1.2 \times 10^5 - 0.504) \text{ Am}^{-1} \\&= 0.151 \text{ T}\end{aligned}$$

The relative permeability

$$\mu_r = \frac{\mathbf{M}}{\mathbf{H}} + 1$$

$$\mu_r = \frac{-0.504 \text{ Am}^{-1}}{1.2 \times 10^5 \text{ Am}^{-1}} + 1 = 0.99$$

Example 14.6 The susceptibility of a diamagnetic material is 5.6×10^{-6} . Assuming a body-centered cubic structure with lattice constant 2.53 \AA , find the mean radius of its atom, if only one electron per atom is contributing to diamagnetisation.

Solution: Since there are 2 atoms in the unit cell of a bcc structure (Section 6.5), the number of electrons per unit volume is given by

$$\begin{aligned} N &= \frac{2}{a^3} = \frac{2}{(2.55 \times 10^{-10} \text{ m})^3} \\ &= 12.06 \times 10^{28} \text{ m}^{-3} \end{aligned}$$

From Eq. (14.21), we get

$$\begin{aligned} \langle r^2 \rangle &= \frac{\chi 6m}{\mu_0 Ze^2 N} \\ r &= \left[\frac{(5.6 \times 10^{-6}) \times 6 (9.1 \times 10^{-31} \text{ kg})}{(4\pi \times 10^{-7} \text{ TmA}^{-1}) \times 1 \times (1.6 \times 10^{-19} \text{ C})^2 (12.06 \times 10^{28} \text{ m}^{-3})} \right]^{\frac{1}{2}} \\ &= 0.888 \times 10^{-10} \text{ m} = 0.888 \text{ \AA} \end{aligned}$$

Example 14.7 The molecular weight of a paramagnetic salt is 168.5 and its density is 4370 kg/m^3 at 27°C . Calculate its susceptibility and the magnetization produced in it in a field of $2 \times 10^5 \text{ A/m}$. Assume that the contribution to paramagnetism is two Bohr magnetons per molecule. Avogadro number = $6.02 \times 10^{26}/\text{k mol}$.

Solution: $N = \frac{\rho N_A}{M}$, N_A = Avogadro number

$$= \frac{(4370 \text{ kg m}^{-3})(6.02 \times 10^{26} / \text{k mol})}{168.5 \text{ kg/k mol}} = 1.56 \times 10^{28} / \text{m}^3$$

$$\begin{aligned} \chi_{para} &= \frac{\mu_0 N \mu}{3kT} = \frac{\mu_0 N p^2 \mu_B^2}{3kT} \\ &= \frac{(4\pi \times 10^{-7} \text{ TmA}^{-1})(1.56 \times 10^{28} / \text{m}^3) \times 4 \times (9.27 \times 10^{-24} \text{ Am}^2)^2}{3(1.38 \times 10^{-23} \text{ JK}^{-1}) 300 \text{ K}} \\ &= 5.43 \times 10^{-4} \end{aligned}$$

REVIEW QUESTIONS

- 14.1** Define magnetic induction (**B**) and magnetization (**M**). State the relation connecting the two.
- 14.2** Define magnetization (**M**) and susceptibility (χ). How does χ vary with temperature for diamagnetic and paramagnetic materials?
- 14.3** Explain the origin of atomic magnetism.
- 14.4** With the necessary expressions explain the magnetic moment of an electron due to its orbital and spin motions.
- 14.5** Outline the classification of magnetic materials.
- 14.6** How are magnetic materials classified from the atomic point of view?
- 14.7** What is diamagnetism? Give examples.
- 14.8** Explain Pauli paramagnetism with examples.
- 14.9** What is Curie paramagnetism? Give two examples.
- 14.10** Distinguish between dia, para, and ferromagnetisms.
- 14.11** Explain the origin of exchange interaction.
- 14.12** State and explain the Curie–Weiss law. How does it differ from Curie's law?
- 14.13** What is saturation magnetization? How does it vary with temperature?
- 14.14** Explain the characteristic features of ferromagnetism.
- 14.15** What are ferromagnetic domains? What is a Bloch wall?
- 14.16** What is hysteresis? Explain the significance of the MH product.
- 14.17** Explain antiferromagnetism with an example.
- 14.18** What are Curie and Neel temperatures?
- 14.19** Distinguish between hard and soft magnetic materials. Give two examples for each.
- 14.20** Explain ferrimagnetism.
- 14.21** What are ferrites? Explain some of their uses.
- 14.22** Represent the spin states of the ferrite Fe_3O_4 .
- 14.23** Explain how antiferromagnetism is explained on the basis of the two sublattice model.
- 14.24** Outline the atomic origin of magnetic moment of atoms. How are magnetic materials classified from the atomic point of view?
- 14.25** Describe hysteresis loop, clearly explaining the ideas regarding coercive field and remanent magnetisation. What are hard and soft magnetic materials? Explain some of their uses.
- 14.26** Discuss with examples the spin arrangements in ferromagnetic, antiferromagnetic, and ferrimagnetic materials. Write a note on ferrites.
- 14.27** How is ferromagnetism explained on the basis of exchange interaction? Give a brief account of Weiss theory of ferromagnetism.

PROBLEMS

1. A paramagnetic substance is placed in an external magnetic field of 10^6 Am^{-1} . Calculate the average value of its magnetic moment per dipole at temperature 27°C .
2. Given the magnetic susceptibility of copper is 7.8×10^{-4} , calculate the flux density and magnetization in copper when it is subjected to a magnetic field of 10^5 Am^{-1} .
3. The paramagnetic susceptibility of copper is 8×10^{-4} . Calculate the dipole moment of a bar of this material of length 1 m and area of cross section 1 cm^2 when placed in a field of $4 \times 10^4 \text{ Am}^{-1}$.
4. Assume that iron atoms have magnetic moment of two Bohr magnetons. Calculate the Curie constant, if its density is 7150 kg/m^3 and atomic weight is 55.84.
5. The magnetic field strength of Zn is 10^4 Am^{-1} . Calculate the magnetization and flux density if its susceptibility is -2×10^{-5} .
6. The density of a paramagnetic sample is 8920 kg/m^3 and its atomic weight is 63.5. The magnetic moment of each atom is one Bohr magneton. Compute its paramagnetic susceptibility at 300 K. Also, calculate the magnetization produced in it in a magnetic field of 10^5 Am^{-1} .

C H A P T E R

15

Thermal Properties

The study of heat capacity of solids at low temperature led to the development of new theories which helped to understand the dynamical behaviour of atoms in solids better. The study mainly considered the lattice and electronic heat capacities. The lattice specific heat refers to the way in which the atoms in the solid behaved with the absorbed heat energy. The electronic specific heat is that due to free electrons in orbitals within an energy range $\approx kT$ near the Fermi level. In this chapter we shall discuss the salient features of the theories of Einstein and Debye with regard to lattice heat capacity of solids. Heat conduction in solids and lattice vibrations are also discussed.

15.1 SPECIFIC HEAT

The ratio of the quantity of heat ΔQ supplied to a body to its corresponding temperature rise ΔT is called the heat capacity C of the body. This is given by

$$C = \frac{\Delta Q}{\Delta T} \quad (15.1)$$

The heat capacity per unit mass, called **specific heat**, is characteristic of the material and is equal to the quantity of heat supplied to unit mass of the material to raise its temperature by one degree. Depending on the condition under which the heating takes place we have specific heat at constant volume (C_V) and specific heat at constant pressure (C_P), which in most cases is around 1 to 2% higher than C_V . By specific heat, we usually mean the specific heat at constant volume, which is more fundamental than heat capacity at constant pressure, since at constant volume, ΔQ is simply the change in internal energy E .

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V$$

According to the classical model, each atom in a solid vibrates about its mean position like a harmonic oscillator. The average energy associated with a harmonic oscillator is kT , k being the Boltzmann constant. A solid containing N atoms is equivalent to $3N$ harmonic oscillators. Hence, the energy of the solid containing N atoms is given by

$$E = 3NkT = 3RT \quad (15.2)$$

where R is the universal gas constant. Hence,

$$C_V = \frac{dE}{dT} = 3R \quad (15.3)$$

That is, the specific heat at constant volume is a constant independent of temperature. This equation, also referred to as **Dulong and Petit's law**, is in agreement with experimental results at high temperatures. But at low temperatures the specific heat is found to decrease with temperature and finally reaches zero at 0 K. This discrepancy was solved by Einstein to some extent.

15.2 EINSTEIN MODEL

In 1906, Einstein proposed a new theory based on the quantum theory for the lattice heat capacity of solids. He assumed that the atoms in a solid vibrate independently with the same frequency v . The quantized energy values of these oscillators, as per quantum theory are given by

$$E_n = nhv = n\hbar\omega, \quad n = 1, 2, 3, \dots, \infty \quad (15.4)$$

The probability that an energy state E_n is occupied is given by the Boltzmann factor $\exp\left(-\frac{E_n}{kT}\right)$.

The average energy \bar{E} of an atomic oscillator is then given by

$$\bar{E} = \frac{\sum_{n=0}^{\infty} E_n N_n}{\sum_{n=0}^{\infty} N_n} = \frac{\sum_{n=0}^{\infty} E_n e^{-E_n/kT}}{\sum_{n=0}^{\infty} e^{-E_n/kT}}$$

Substituting the value of E_n from Eq. (15.4), we get

$$\begin{aligned} \bar{E} &= \frac{\sum_{n=0}^{\infty} nhv e^{-\frac{nhv}{kT}}}{\sum_{n=0}^{\infty} e^{-\frac{nhv}{kT}}} = \frac{hve^{-\frac{hv}{kT}} + 2hve^{-\frac{2hv}{kT}} + \dots}{1 + e^{-\frac{hv}{kT}} + e^{-\frac{2hv}{kT}} + \dots} \end{aligned}$$

Writing $e^{-\frac{hv}{kT}} = x$, we get

$$\begin{aligned}\bar{E} &= \frac{hv(x(1+2x+3x^2+\dots))}{1+x+x^2+\dots} = \frac{hv(x(1-x)^{-2})}{(1-x)^{-1}} \\ &= \frac{hv x}{1-x} = \frac{hv e^{-\frac{hv}{kT}}}{1-e^{-\frac{hv}{kT}}} \\ &= \frac{hv}{e^{\frac{hv}{kT}} - 1}\end{aligned}\quad (15.5)$$

This represents the average energy of an atomic oscillator. For $3N$ oscillators (for a solid with N atoms and three directions for each), the energy is given by

$$\begin{aligned}E &= 3N\bar{E} = \frac{3Nhv}{e^{\frac{hv}{kT}} - 1} \\ C_V &= \frac{dE}{dT} = 3Nk \left(\frac{hv}{kT}\right)^2 \frac{e^{\frac{hv}{kT}}}{\left(e^{\frac{hv}{kT}} - 1\right)^2}\end{aligned}\quad (15.6)$$

which is Einstein's expression for lattice specific heat. It is convenient to introduce a characteristic temperature, called **Einstein temperature**, defined by

$$\theta_E = \frac{hv}{k}\quad (15.7)$$

$$C_V = 3R \left(\frac{\theta_E}{T}\right)^2 \frac{e^{\frac{\theta_E}{T}}}{\left(e^{\frac{\theta_E}{T}} - 1\right)^2}\quad (15.8)$$

15.2.1 High Temperature Limit

At high temperatures, $T \gg \theta_E$, and therefore,

$$\begin{aligned}e^{\frac{\theta_E}{T}} &\approx 1 + \frac{\theta_E}{T} \approx 1 \\ e^{\frac{\theta_E}{T}} - 1 &\approx \frac{\theta_E}{T}\end{aligned}$$

Hence,

$$C_V = 3R \left(\frac{\theta_E}{T} \right)^2 \frac{1}{\left(\frac{\theta_E}{T} \right)^2} = 3R \quad (15.9)$$

which is the same as the classical result.

15.2.2 Low Temperature Limit

At low temperatures, $T \ll \theta_E$, and therefore,

$$e^{\frac{\theta_E}{T}} - 1 \approx e^{\frac{\theta_E}{T}}$$

Hence,

$$\begin{aligned} C_V &= 3R \left(\frac{\theta_E}{T} \right)^2 \frac{1}{e^{\frac{\theta_E}{T}}} \\ &= 3R \left(\frac{\theta_E}{T} \right)^2 e^{-\frac{\theta_E}{T}} \end{aligned} \quad (15.10)$$

At low temperatures, the exponential term is more important than the term $\left(\frac{\theta_E}{T} \right)^2$. Experiments reveal that C_V approaches zero as T^3 . However, the decrease predicted by Eq. (15.10), is very rapid than that obtained by experimental observations (Fig. 15.1). Thus, the Einstein's model explains the observed specific heat of solids at high temperatures, but fails at very low temperatures. This was resolved by Debye in his model.

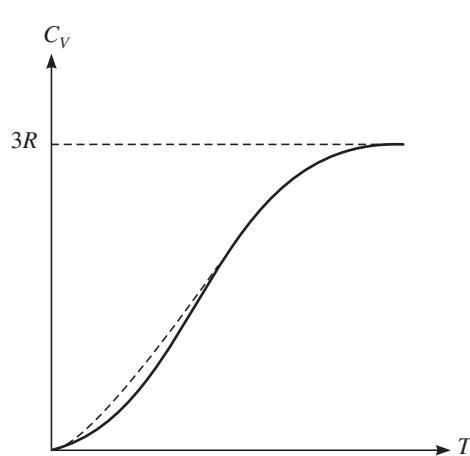


Fig. 15.1 Specific heat C_V as a function of temperature T for a typical solid. The solid line is as per the Einstein model, Eq. (15.10). The dotted line shows the experimental results.

15.3 DEBYE MODEL

In the Einstein model, he assumed that all the atoms in the solid vibrate independently with a constant frequency. By contrast Debye treated the atoms of the crystal as coupled harmonic oscillators having a range of frequencies unlike the constant frequency of Einstein. Let us consider the vibrational motion of the crystal as a whole rather than the vibrational motion of a single atom. Like Einstein, Debye assumed that the allowed energy of these vibrations are given by

$$E_n = nh\nu = n\hbar\omega, \quad n = 0, 1, 2, \dots$$

and the average energy of a vibrator

$$\bar{E} = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1} \quad (15.11)$$

Since Debye assumed a continuous range of frequencies, one should have an expression for a quantity $D(\nu)$ which represents the number of modes per unit frequency range at the frequency ν . This quantity is often referred to as the **density of states**. Consequently, the quantity, $D(\nu) d\nu$ represents the number of modes in the frequency range ν and $(\nu + d\nu)$. The contribution made by these modes to the internal energy of the crystal is then given by

$$dE = \bar{E}(\nu) D(\nu) d\nu \quad (15.12)$$

where $\bar{E}(\nu)$ is the average energy density of a mode of frequency ν . The total energy is then

$$E = \int dE = \int \bar{E}(\nu) D(\nu) d\nu \quad (15.13)$$

from which we can calculate C_V . Therefore, for calculating the specific heat C_V , essentially one requires the value of $D(\nu)$. Assuming the same velocity for the transverse and longitudinal modes, it can be proved that

$$D(\nu) = \frac{12\pi V \nu^2}{\nu^3} \quad (15.14)$$

where V is the volume of the solid and v is the constant velocity of sound in the solid.

The upper limit for the range of frequencies cannot be infinite since it leads to an unphysical situation. Therefore, Debye made a cut off for the upper frequency such that the total number of modes of vibration becomes $3N$, since N atoms will have $3N$ modes of vibration. This procedure led Debye to a maximum frequency ν_D , called **Debye frequency**. Since D_v represents the number of modes per unit frequency range, we have

$$\int_0^{\nu_D} D(\nu) d\nu = 3N \quad \text{or} \quad \frac{12\pi V}{v^3} \int_0^{\nu_D} \nu^2 d\nu = 3N$$

$$\frac{12\pi V}{v^3} \frac{\nu_D^3}{3} = 3N$$

$$v_D = \left(\frac{9N}{12\pi V} \right)^{\frac{1}{3}} v \quad (15.15)$$

This gives the maximum frequency of the mode of vibration that can propagate in the solid.

Substituting the value of $D(v)$ from Eq. (15.14) and \bar{E} from Eq. (15.11) in Eq. (15.13) to find an expression for energy, we get

$$\begin{aligned} E &= \int_0^{v_D} \frac{hv}{e^{\frac{hv}{kT}} - 1} \frac{12\pi V v^2}{v^3} dv \\ &= \frac{12\pi h V}{v^3} \int_0^{v_D} \frac{v^3 dv}{e^{\frac{hv}{kT}} - 1} \end{aligned} \quad (15.16)$$

Writing

$$\frac{hv}{kT} = x \quad \text{and} \quad \frac{hv_D}{kT} = x_D \quad (15.17)$$

and using Eq. (15.15), the expression for energy, Eq. (15.16), becomes

$$\begin{aligned} E &= \frac{12\pi h V}{v^3} \left(\frac{kT}{h} \right)^4 \int \frac{x^3 dx}{e^x - 1} \\ &= 9N \left(\frac{kT}{hv_D} \right)^3 kT \int_0^{x_D} \frac{x^3 dx}{e^x - 1} \end{aligned} \quad (15.18)$$

Following the expression $E = kT$, we define a **Debye temperature** θ_D by

$$E = k\theta_D = hv_D \quad \text{or} \quad \theta_D = \frac{hv_D}{k} \quad (15.19)$$

Now using Eq. (15.17), we obtain

$$\theta_D = \frac{hv_D}{k} = x_D T \quad (15.20)$$

In terms of θ_D , Eq. (15.18) takes the form

$$E = 9NkT \left(\frac{T}{\theta_D} \right)^3 \int_0^{x_D} \frac{x^3 dx}{e^x - 1} \quad (15.21)$$

Case (i) $T \gg \theta_D$: For very high temperatures x is very small, and therefore, $e^x \approx 1 + x$. Eq. (15.21) takes the form

$$E = 9NkT \left(\frac{T}{\theta_D} \right)^3 \int_0^{\theta_D} x^2 dx = 9NkT \left(\frac{T}{\theta_D} \right)^3 \left(\frac{\theta_D}{T} \right)^3 \frac{1}{3} = 3RT \quad (15.22)$$

$$C_V = \frac{\partial E}{\partial T} = 3R$$

which is the same as Dulong and Petit's law.

Case (ii) $T \ll \theta_D$: For very low temperatures, the upper limit is infinity and using a table of integrals

$$\int_0^{\infty} \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}$$

Consequently,

$$E = 9NkT \left(\frac{T}{\theta_D} \right)^3 \frac{\pi^4}{15} = \frac{3}{5} \pi^4 NkT \left(\frac{T}{\theta_D} \right)^3 \quad (15.23)$$

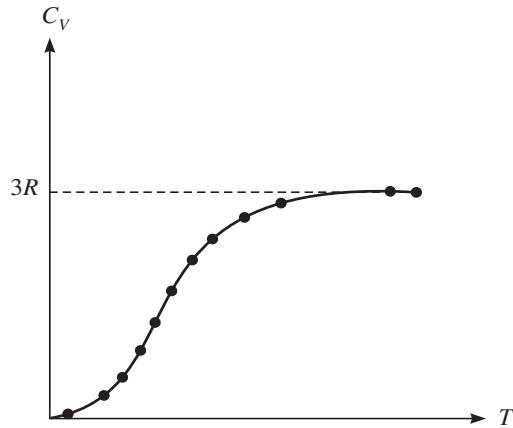


Fig. 15.2 Specific heat C_V as a function of temperature T .

The solid line is as per Debye model, Eq. (15.24) becomes

$$C_V = \frac{12}{5} \pi^4 Nk \left(\frac{T}{\theta_D} \right)^3 \quad (15.24)$$

Thus, the lattice specific heat C_V varies as T^3 which is Debye's T^3 law. Agreement between theory and experiment is quite good, as shown in Fig. 15.2.

15.4 HEAT CAPACITY OF ELECTRONS

We shall now consider the effect of the valence electrons in a metal on the heat capacity. The classical theory failed to explain the experimentally observed results and the problem was solved with the help of the Fermi-Dirac distribution (Section 10.9). Accordingly, the electronic heat capacity is given by Eq. (10.49):

$$C_e = \frac{\pi^2 k^2 N T}{2E_F} \quad (15.25)$$

where E_F is the Fermi energy, k is the Boltzmann constant, T is the absolute temperature, and N is the total number of valence electrons. For a monovalent metal, the number of valence electrons in one mole is simply the Avogadro's number (N_A). For a divalent metal, the number per mole is $2N_A$, and so on.

15.5 LATTICE VIBRATIONS

The arrangement of atoms in a crystal lattice is periodic. The atoms are connected with each other by a force just like two particles connected by an elastic spring. Therefore, the motion of every atom would be coupled. In other words, the motion of every atom is shared by all the atoms in the crystal and the crystal would vibrate as a whole. We shall now investigate some of the characteristics of these elastic vibrational motion, called **lattice vibrations**, of the crystal lattice.

15.5.1 One-dimensional Monoatomic Lattice Vibrations

Born, Von Karmon and Blackman in 1912 proposed an one dimensional model of the solid in which the atoms are bonded to each other, and studied their elastic vibrations. When the lattice is in equilibrium, each atom will be exactly at its lattice site. When the lattice begins to vibrate, each atom is displaced from the lattice site as shown in Fig. 15.3. As mentioned, the motion of the atoms would be coupled, and therefore we have to consider the motion of the entire lattice.

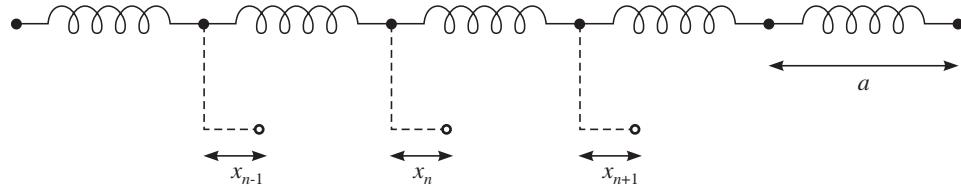


Fig. 15.3 One dimensional chain of atoms.

x_{n-1} , x_n , and x_{n+1} are the displacements of the $(n - 1)$, n , and $(n + 1)$ th atoms. When a wave propagates through the lattice, each wave vector has 3 modes, one of longitudinal and two of transversal polarizations. The equations of motions are then written on the basis of the following assumptions:

- (i) The force acting on an atom is only due to the nearest neighbours.
- (ii) The forces are linear. That is, the forces are proportional to the relative displacement of these atoms. This is known as the **harmonic approximation**.

Solution of the equations of motion gives the following **dispersion relations** for the one dimensional lattice

$$\omega^2 = \frac{4\alpha}{m} \sin^2 \frac{ka}{2} = \omega_{\max}^2 \sin^2 \frac{ka}{2} \quad (15.26)$$

where α is the force constant, m is the mass of the atom, k is the wave vector of the wave and ω is the angular frequency. Figure 15.4 shows the variation of ω with k .

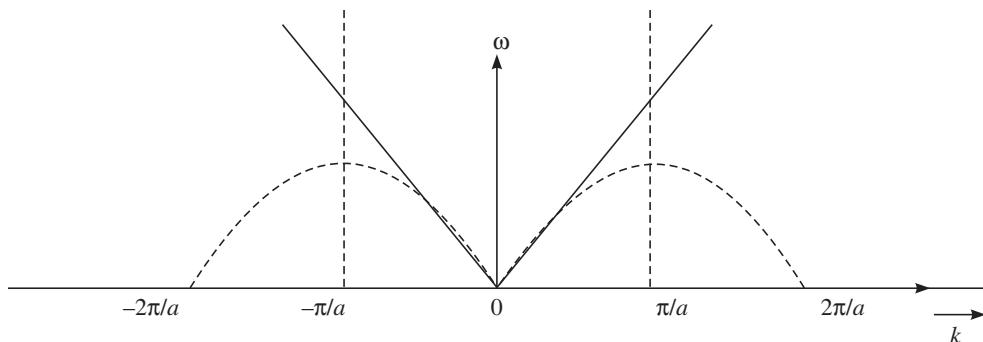


Figure 15.4 Dispersion curve for monoatomic lattice.

15.6 LATTICE WITH TWO ATOMS PER PRIMITIVE CELL

In a diatomic case like $NaCl$, the atoms of two kinds of masses M and m are arranged alternatively as shown in Fig. 15.5.

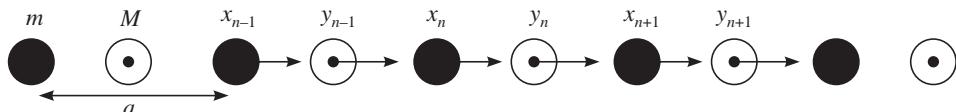


Fig. 15.5 One dimensional diatomic linear lattice with atomic masses m and M .

Assuming that only the nearest neighbour interactions are significant, the equations of motion are formed which led to the following equation:

$$mM\omega^4 - 2\alpha(m + M)\omega^2 + 2\alpha^2(1 - \cos ka) = 0 \quad (15.27)$$

where α is the force constant, k is the wave vector, and ω is the angular frequency. For each polarization in a given propagation direction, ω versus k dispersion relation develops two branches called the **acoustical** and **optical branches**. The corresponding phonons are called **acoustic** and **optical phonons**, respectively. The acoustic and optical branches of the dispersion relations are illustrated in Fig. 15.6.

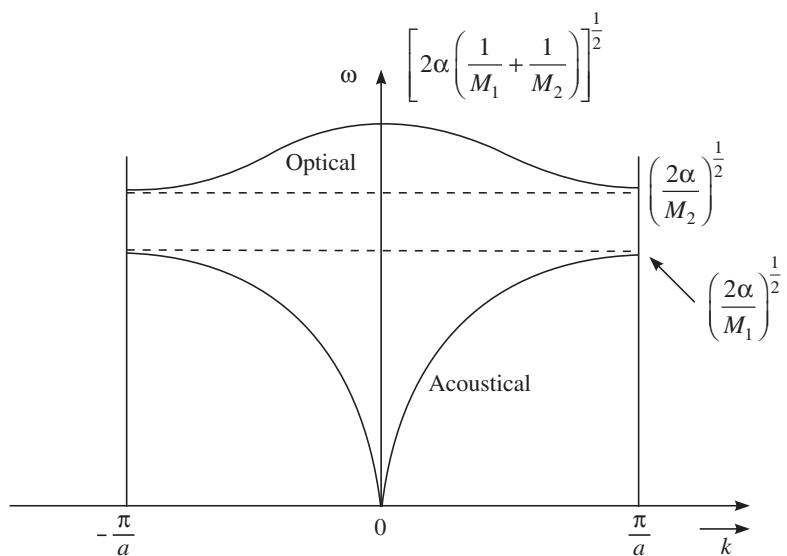


Fig. 15.6 The two branches of the dispersion curve of a linear diatomic lattice.

For the optical branch, the two atoms vibrate in opposite directions, keeping their centre of mass fixed. If the two atoms have opposite charges as in an ionic crystal, the electric field associated with the light wave can excite this type of motion, hence the name optical branch. Another reason for this name is that its frequency lies in the infrared region ($\approx 3 \times 10^{13}$ Hz). For the acoustical branch, the two atoms will be moving in phase, and therefore the whole lattice is oscillating as a single rigid one. This vibration can be excited by a sound wave, hence the name acoustical branch. The range between the top of the acoustical branch and the bottom of the optical branch is forbidden. The lattice does not transmit a wave of these frequencies, thus acting as a band pass filter.

15.7 PHONONS

At a particular temperature, the atoms in the lattice of a solid vibrate about their equilibrium position. The wave thus produced in a solid is characterised by a wavevector k , propagation velocity v , and wavelength λ . The lattice vibrational energy is quantized and the quanta of energy is called **phonon**. Different kinds of lattice vibration are possible. Sound waves are acoustic phonons and the optical branch is due to optical phonons. Thermally excited phonons create thermal vibrations in the crystal. The phonon spectra have frequencies ranging from $10^4 - 10^{13}$ Hz. A phonon of wave vector k interacts with particles like electrons, photons, neutrons, etc. and fields as if it has a momentum $\hbar k$. Phonon momentum is often referred to as **crystal momentum**.

Phonons have zero spin and obey Bose-Einstein statistics. Hence, phonons are bosons and the average number of phonons in a mode is given by

$$N = \frac{1}{e^{\frac{\hbar\omega}{kT}} - 1} \quad (15.28)$$

This equation tells that phonons can be created by raising the temperature and hence the number of phonons in a system is not conserved.

15.8 THERMAL CONDUCTIVITY

To explain the high electrical and thermal conductivities of metals, Sommerfeld proposed the model known as **free electron model**. It is assumed that the valence electrons in the metals are not localized and are free to move around. A conducting metal can be thought of as a collection of non-interacting electrons that are free to roam about the specimen, obeying Fermi-Dirac statistics.

Let us consider a metallic rod whose ends are at different temperatures as in Fig. 15.7. The amount of heat energy flowing through unit area in unit time from the hot (left end) to the cold end is given by

$$Q = K \frac{dT}{dx} \quad (15.29)$$

where K is the **coefficient of thermal conductivity** and $\frac{dT}{dx}$ is the temperature gradient. Let

AA' , BB' and CC' be three sections of the rod perpendicular to the flow of heat. Let the sections be separated by a distance λ , the mean free path of the electrons. Let the temperature and heat energy experienced by the middle section BB' be T and E . Then the temperature and energy

at the sections AA' and CC' will be $T + \left(\frac{dT}{dx}\right)\lambda$, $E + \left(\frac{dE}{dx}\right)\lambda$ and $T - \left(\frac{dT}{dx}\right)\lambda$, $E - \left(\frac{dE}{dx}\right)\lambda$,

respectively. If n is the number of electrons per unit volume moving with an average velocity v , according to kinetic theory, the net amount of energy flowing through the plane BB' is given by

$$Q = \frac{1}{3} nv\lambda \frac{dE}{dx} \quad (15.30)$$

This must be equal to Q in Eq. (15.29). That is,

$$\begin{aligned} K \frac{dT}{dx} &= \frac{1}{3} nv\lambda \frac{dE}{dx} \\ &= \frac{1}{3} nv\lambda \left(\frac{dE}{dT}\right) \left(\frac{dT}{dx}\right) \end{aligned} \quad (15.31)$$

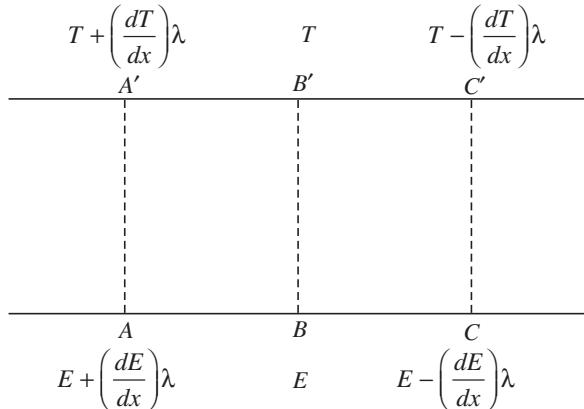


Fig. 15.7 Heat conduction across the planes \$AA'\$, \$BB'\$ and \$CC'\$.

As we are concerned with the energy transfer accomplished by the electrons, $\left(\frac{dE}{dT}\right)$ is the electronic heat capacity C_e of a particle. Hence,

$$\begin{aligned} K &= \frac{1}{3} n v c_e \lambda \\ &= \frac{1}{3} v C_e \lambda \end{aligned} \quad (15.32)$$

where C_e is the electronic specific heat per unit volume, v is the velocity of electrons, and λ is the mean free path. The electronic specific heat C_e is given by Eq. (15.25). Now the *rms* velocity v changes to v_F , the Fermi speed, and λ change to λ_F . Substituting the values in Eq. (15.32), we get

$$K = \frac{\pi^2 N k^2 T v_F \lambda_F}{6 E_F} \quad (15.33)$$

Since

$$E_F = \frac{1}{2} m v_F^2 \quad \text{or} \quad 2 E_F = m v_F^2 \quad (15.34)$$

and writing $\frac{\lambda_F}{v_F} = \tau_F$

$$K = \frac{\pi^2 N k^2 T \tau_F}{3m} \quad (15.35)$$

In a solid, both electrons and phonons are carriers of thermal energy. In pure metals, the electronic contribution exceeds that of the phonons at all temperatures. However, in alloys and impure metals, the phonons make appreciable contribution. This is because of the reduction in mean free path of electrons due to collisions with impurities.

SOLVED EXAMPLES

Example 15.1 The highest possible frequency allowed to pass through aluminium is 64×10^{11} Hz. Calculate its Debye temperature.

Solution: Debye temperature is given by

$$\theta_D = \frac{h\nu_D}{k} = \frac{(6.626 \times 10^{-34} \text{ Js})(64 \times 10^{11} \text{ s}^{-1})}{1.38 \times 10^{-23} \text{ JK}^{-1}}$$

$$= 307.3 \text{ K}$$

Example 15.2 In the Debye model, the highest vibrational frequency of carbon is 40.5×10^{12} Hz. Calculate the lattice specific heat of carbon at 30 K.

$$\theta_D = \frac{h\nu_D}{k} = \frac{(6.626 \times 10^{-34} \text{ Js})(40.5 \times 10^{12} \text{ s}^{-1})}{1.38 \times 10^{-23} \text{ JK}^{-1}} = 1944.59 \text{ K}$$

The lattice specific heat is given by

$$C_l = \frac{12}{5} \pi^4 N k \left(\frac{T}{\theta_D} \right)^3$$

$$= \frac{12}{5} \pi^4 (6.02 \times 10^{26} \text{ k mol}^{-1}) (1.38 \times 10^{-23} \text{ JK}^{-1}) \left(\frac{30 \text{ K}}{1944.59 \text{ K}} \right)^3$$

$$= 7.12 \text{ J k mol}^{-1} \text{ K}^{-1}$$

Example 15.3 Show that the Einstein frequency for copper falls in the infrared region if its Einstein temperature is 1990 K.

$$\nu_E = \frac{k\theta_E}{h} = \frac{(1.38 \times 10^{-23} \text{ JK}^{-1})(1990 \text{ K})}{(6.626 \times 10^{-34} \text{ Js})}$$

$$= 4.14 \times 10^{13} \text{ Hz}$$

This frequency falls in the near infrared region.

Example 15.4 Calculate the electronic and lattice heat capacities at 0.05 K for copper. The Fermi energy of copper is 7 eV; its Debye temperature is 348 K.

Solution: The electronic heat capacity of copper is given by

$$\begin{aligned} C_e &= \frac{\pi^2 N k^2 T}{2E_F} \\ &= \frac{\pi^2 (6.02 \times 10^{23} \text{ mol}^{-1}) (1.38 \times 10^{-23} \text{ JK}^{-1})^2 0.05 \text{ K}}{2 \times 7 \text{ eV} \left(1.6 \times 10^{-19} \frac{\text{J}}{\text{eV}} \right)} \\ &= 2.53 \times 10^{-5} \text{ J mol}^{-1} \text{ K}^{-1} \end{aligned}$$

Lattice heat capacity of copper (Eq. 15.24) is:

$$\begin{aligned} C_V &= \frac{12}{5} \pi^4 N k \left(\frac{T}{\theta_D} \right)^3 \\ &= \frac{12}{5} \pi^4 (6.02 \times 10^{23} \text{ mol}^{-1}) (1.38 \times 10^{-23} \text{ JK}^{-1}) \left(\frac{0.05 \text{ K}}{348 \text{ K}} \right)^3 \\ &= 3.24 \times 10^{-5} \text{ J mol}^{-1} \text{ K}^{-1} \end{aligned}$$

Example 15.5 Show that the heat capacity, when the temperature equals the Einstein temperature, is equal to 0.92 of the classical value.

Solution: When the temperature T is equal to the Einstein temperature θ_E , from Eq. (15.8), we have

$$C_V = 3R \frac{e^1}{(e^1 - 1)^2} = 0.92 \times 3R$$

Example 15.6 Calculate the molar electronic heat capacity of zinc at 300 K, assuming that there are 2 valence electrons in zinc. The Fermi energy of zinc = 9.38 eV.

Solution: From Eq. (15.25), the electronic heat capacity is obtained as

$$C_e = \frac{\pi^2 k^2 N T}{2E_F}$$

For 1 mole of zinc, $N = 2$ Avogadro's number

$$\begin{aligned} &= 2 \times 6.02 \times 10^{23} / \text{mol} \\ C_e &= \frac{\pi^2}{2} \frac{(1.38 \times 10^{-23} \text{ JK}^{-1})^2 (2 \times 6.02 \times 10^{23} / \text{mol}) 300 \text{ K}}{(9.38 \text{ eV}) (1.6 \times 10^{-19} \text{ J/eV})} \\ &= 0.226 \text{ J mol}^{-1} \text{ K}^{-1} \end{aligned}$$

REVIEW QUESTIONS

- 15.1** State Dulong and Petit's law. How did Einstein explain the departure of specific heat of solids from this law?
- 15.2** Compare the assumptions and results of Einstein and Debye theories.
- 15.3** Define Debye temperature. How does it differ from Einstein temperature?
- 15.4** Distinguish between lattice heat capacity and electronic heat capacity.
- 15.5** What are acoustical and optical branches in lattice vibrations? And why are they named likewise?
- 15.6** What are lattice vibrations and phonons?
- 15.7** Explain the basic ideas behind the Einstein model for the specific heat of solids. Hence derive an expression for the same.
- 15.8** Explain the Debye's model for the specific heat of a solid. On the basis of this model derive an expression for the lattice heat capacity of a solid.
- 15.9** What are free electrons? Derive an expression for the coefficient of thermal conductivity of a metal.

PROBLEMS

1. Calculate the highest possible frequency that copper allows through it, if its Debye temperature is 350 K.
2. Calculate the specific heat of lead at 10 K if its Debye temperature is 105 K. Also determine the highest frequency that the sample allows to propagate through it.
3. At low temperatures, the Debye temperatures of solids A and B are 300 K and 250 K, respectively. The lattice specific heat of A at 5 K is $0.05 \text{ Jmol}^{-1}\text{K}^{-1}$. Estimate the heat capacity of B at 5 K and that of A at 2 K.
4. Find the temperature at which the lattice specific heat and the electronic specific heat of copper are the same. The Fermi energy and θ_D for copper are 7 eV and 348 K, respectively.
5. Calculate the molar electronic heat capacity of aluminium at 300 K, assuming that there are 3 valence electrons in it. The Fermi energy of aluminium = 11.58 eV.
6. Assuming that the ionic contribution to heat capacity is $3 R$, find the ratio of the electronic and ionic contributions to the heat capacity of a monovalent metal.

C H A P T E R

16

Nanotechnology

The concept of nanotechnology is attributed to Richard Feynman for his talk in 1959 *there is plenty of room at the bottom*. It outlined the prospects for the possibility of manoeuvring things atom by atom. The word nanotechnology was coined by Nario Taniguchi in 1974 to describe machining with tolerances of less than a micron. Feynman's prophecy became a reality with the advent of scanning tunnelling microscope in 1981. The prophecy reached a climax with the publication of the book *Engines of Creation, The Coming Age of Nanotechnology* by Drexler, in 1990. In this chapter, we shall discuss some of the basics of nanotechnology which ensures a bright exciting future to mankind.

16.1 INTRODUCTION

Any material that contains grains or clusters of the order of 1 to 100 nm or layers or filaments of that dimension are considered as **nanomaterials**. When a given volume of matter is divided into a number of small pieces, the total surface area of the pieces is much greater than the area of the given volume of matter. That is, nanoparticles have a much greater surface area per given volume of bulk matter.

Let us consider a row of atoms arranged on a surface. One side of the row of atoms makes contact with the surface and the other side is not in contact with anything. Such atoms are completely different from those that are fully buried in the bulk matter, leading to completely different properties. That is, owing to the small grain size and the high surface to volume ratio, these materials are expected to show enhanced mechanical, magnetic, optical, and chemical properties. Some materials are exceptionally strong, hard, and ductile at high temperatures. Still they are chemically very active, as the number of surface atoms or molecules are very large.

During synthesis, the nanomaterials have the capability to be sintered at much lower temperatures than conventional powders. The enhanced properties of nanomaterials have already been exploited for a variety of structural and non-structural applications.

The development of semiconductor nanoclusters is another area of interest. These nanoclusters are often referred to as **quantum dots**, **nanocrystals**, and **Q-particles**. The quantum dots can be developed to emit and absorb a desired wavelength of light by changing the particle dimension. This feature allows the construction of finely tunable semiconductor laser.

16.2 SYNTHESIS OF NANOMATERIALS—GENERAL IDEAS

Much progress has been achieved during the last 10 years in the synthesis of nanomaterials with controlled size, shape, and crystalline structure. However, it is still difficult to generalise the underlying physical or chemical principles behind the existing methods of any arbitrary nanomaterial. We shall now try to specify some of the routes used for their synthesis.

Synthesis of nanoparticles are typically grouped into two categories: the *top down methods* and the *bottom up methods*. The top down method involves milling, chemical methods for breaking specific bonds such as hydrogen bonds, and volatilization of a solid by laser ablation, solar furnace, or some other method followed by condensation of volatized components. The bottom up method involves condensation of atoms or molecular entities in a gas phase or in solution. Depending on the material and the specific application, the particular approach is selected. For example, metal oxide nanoparticles can be prepared by both methods. However, the most common route involves the bottom up approach by precipitation in aqueous solution from metal salts.

The liquid phase colloidal synthesis of nanoparticles is very powerful, convenient, and reproducible. It allows for the resulting nanocrystals to be precisely tuned to the requirements of size, shape, crystalline structure, and composition on the nanometer scale. It also allows them to be dispersed in either an aqueous or a nonaqueous medium. In addition, these nanoparticles can be modified in liquid suspension by treatment with various chemical species for application and use in technical and biological systems. We discuss in brief some of the methods commonly used for the preparation of nanoparticles.

16.3 LASER INDUCED EVAPORATION METHOD

Laser induced evaporation method is a very powerful method for the preparation of nanoparticles of metal atoms. A high intensity laser beam is allowed to fall on a metal rod, causing evaporation of atoms from the surface of the metal. Figure 16.1 illustrates such a device. The evaporated atoms are then swept away by a burst of helium gas and passed through an orifice into a vacuum chamber. The sudden expansion of the gas produces cooling, which causes the formation of clusters of metal atoms. These clusters are used as such or irradiated by UV radiation which makes them into smaller clusters. The resulting nanoparticles are then allowed to escape through an orifice in the chamber where it is used for different purposes. If needed, a skimmer can also be introduced as shown in Fig. 16.1. It is noticed that clusters of 7 to

10 atoms are more likely than other clusters. Gases can also be introduced to study the chemical interaction of the gases with the nanoparticles.

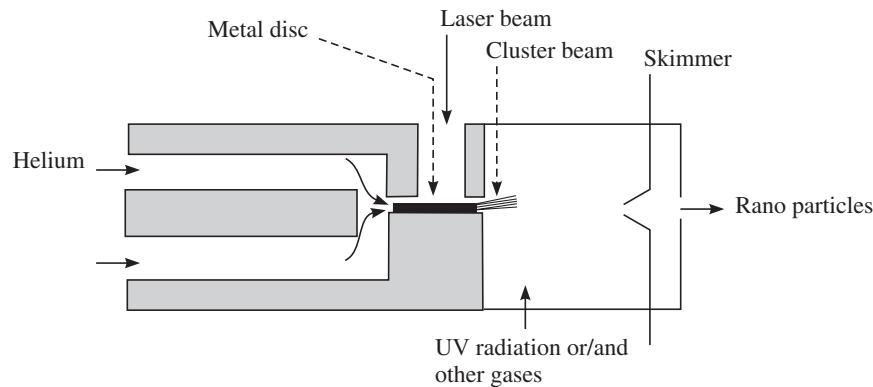


Fig. 16.1 Laser induced evaporation method.

16.4 LASER PULSE METHOD

The laser pulse method is commonly used for the preparation of nanoparticles of silver. Silver nitrate solution and a reducing agent are mixed in a blender as illustrated in Fig. 16.2. In the blender there is a solid disc, which is fixed to a sample holder. When the sample holder rotates, the disc too rotates in the solution. The pulsed laser beam directed towards the disc creates hot spots on the surface of the disc. The silver nitrate and the reducing agent react at these hot spots, resulting in the formation of nanoparticles of silver. The particles can be separated from the solution using a centrifuge. The size of the particles can be controlled by adjusting the rotation speed of the disc and the energy of the laser.

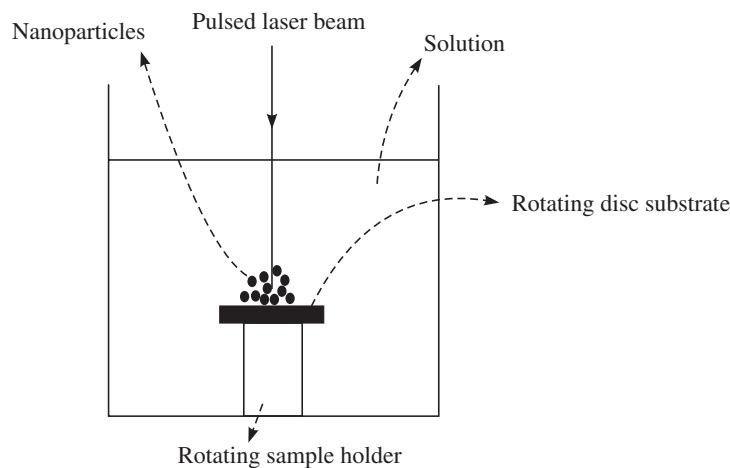


Fig. 16.2 Set up to make silver nanoparticles using a pulsed laser beam.

16.5 METAL NANOCLOUDS – RF PLASMA METHOD

In this method, nanoparticles are synthesised by using a plasma generated by radiofrequency (RF) heating coils. The metal whose nanoparticles are required is taken in a small pestle and kept inside an evacuated chamber. The metal is heated above the evaporation point using high voltage RF coils wound around the evacuated chamber close to the pestle having the metal (Fig. 16.3). Helium gas is then allowed to enter the system, which forms a high temperature plasma in the region of the coils. The metal vapour nucleates on the helium atoms and diffuse upwards to the colder collector rod where the nanoparticles are deposited. The particles are passivated by the introduction of a gas, like oxygen. In the case of aluminium nanoparticles, the oxygen forms a layer of aluminum oxide around the particle.

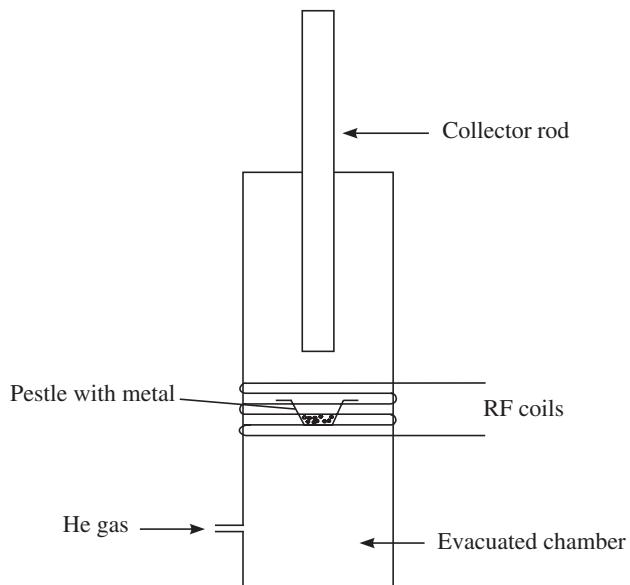


Fig. 16.3 Apparatus for the synthesis of nanoparticles using an RF produced plasma.

16.6 THERMOLYSIS

The process of nanoparticle preparation by thermally decomposing solids consisting of metal cations and molecular anions or metal organic solids is called **thermolysis**. The material, say Li N_3 , is placed in an evacuated quartz tube and heated to 675 K temperature in the experimental arrangement illustrated in Fig. 16.4. At about 645 K, the Li N_3 decomposes into Li and N_2 gas. This can be noticed by an increase in the pressure on the vacuum gauge. In a few minutes, the pressure again drops to the original low value, indicating the removal of all the N_2 gas. The lithium atoms coalesce to form small colloidal metal particles. Particles of 3 – 5 nm size can be made by this method. Passivation can be achieved by introducing proper gases. The heating temperature depends on the material selected for the preparation.

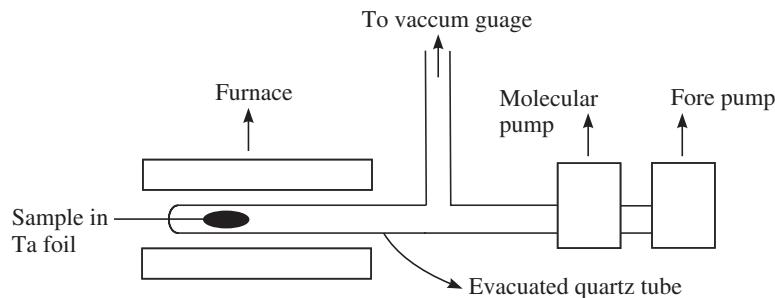


Fig. 16.4 Arrangement for the preparation of metal nanoparticles by thermally decomposing solids.

16.7 CHEMICAL VAPOUR DEPOSITION

In chemical vapour deposition, the material whose nanoparticles are needed is heated to form a gas, and then allowed to deposit as a solid on a surface, usually under vacuum. In addition to direct deposition, deposition by chemical reaction to form a new product that is different from the volatilized product is also possible. If vapours of carbon or oxygen are present with the metal, nanopowders of carbides and oxides can readily be obtained. For the production of pure metal powders, microwaves tuned to the metal excitation frequencies are used to melt and vapourize the materials to produce a plasma at temperatures of the order of 1700–1800 K. The plasma then enters a water cooled reaction column which helps the formation of nanosize particles. Complete conversion is possible by adjusting the residence time of reactants in the plasma state. After the nanoparticles attain appropriate cooling temperature (900–1000 K), they are filtered from the exhaust gas flow and fall into a container. The factors that decide the particle grain size are the metal concentration in the gas phase, flow rate of metal vapour, and temperature.

Chemical vapour deposition can also be used to coat surfaces by keeping the object to be coated in the presence of the chemical vapour. The first layer of atoms or molecules deposited may not form a good coating with the surface. In such a case, the deposited species can act as a template on which materials can grow. As the atoms or molecules deposited can be influenced by their neighbours, often structures of these materials are aligned. Sometimes a site for crystallization may form which may result in the vertical growth of the structures. Usually a material is coated on a surface and then parts of the coating are removed using a strong electron beam or chemicals. Chemical vapour deposition is then used to build structures on these surfaces (Fig. 16.5).

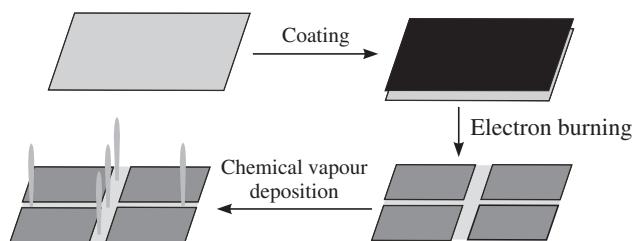


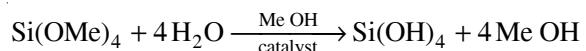
Fig. 16.5 Nanopillars by chemical vapour deposition.

16.8 SOL-GEL METHOD

The sol-gel process is a technique to make metal oxide and non-metal oxide glasses through chemical reactions without the use of high temperature processing. The sol-gel process occurs in four stages as follows:

- (i) Hydrolysis
- (ii) Condensation and polymerization
- (iii) Growth of particles
- (iv) Agglomeration of particles followed by formation of networks

The precursors for the synthesis normally consists of the ions of a metal. Metal alkoxides and alkoxy silanes are most popular because they react readily with water. The commonly used alkoxy silanes are tetra methoxysilane (TMOS) and tetraethoxysilane (TEOS), which form silica gels. The addition of an acid or base catalyst to a solution of an alkoxy silane reagent, say TMOS, water and methanol leads to the hydrolysis of the Si-OMe bonds to form Si-OH functional group. The reaction is given as follows:



The addition of external catalysts speeds up and completes the hydrolysis. Subsequent elimination of water from Si(OH)_4 gives a condensed three dimensional oxide network. On drying, the trapped water and alcohol are removed, the network shrinks, and the condensation process moves on further. The resulting gel matrix is often called **xerogel**.

As hydrolysis and condensation reactions occur concurrently, the products would be monomers, dimers, linear trimers, cyclic trimers, cyclic tetramers, and higher order rings. The characteristics and properties of a particular sol-gel network are related to the following factors:

- (i) Reagent concentrations
- (ii) pH
- (iii) Temperature and time of reaction
- (iv) $\text{H}_2\text{O/Si}$ molar ratio
- (v) Nature and concentration of catalyst
- (vi) Drying

Of these, pH, $\text{H}_2\text{O/Si}$ molar ratio, and temperature are the most important ones.

The gel products have properties ideal for desired applications. Porous ceramic xerogels of high surface area can serve as supports for thin film deposition. Between pH-2 and pH-6, the rate of dimerization and consequent polymerization is almost nil. Further growth occurs by the addition to lower molecular weight species to more highly condensed species and aggregation of the condensed species to form chain and networks. Particle growth stops when the particles reach 2 – 4 nm in diameter.

Mixtures of elements can also undergo hetero-condensation to form mixed element gels. The possibility of having a wide variety of precursors makes the sol-gel method an extremely useful one for the preparation of nanocomposite material.

16.9 CARBON CLUSTERS

The number of electrons in a carbon atom is 6. Its normal ground state configuration is $1s^2 2s^2 2p_x^1 2p_y^1 2p_z^0$, and one expects the carbon atom to form two bonds. However, due to the hybridization of the 2s and 2p orbitals, 4 orbitals of the same shape and energy are available for bonding. Hence, carbon becomes tetravalent. Three types of hybridizations are possible—sp hybridization, sp^2 hybridization, and sp^3 hybridization. These hybridizations result in bond angles of 180° as in C_2H_2 , 120° as in C_2H_4 , and $109^\circ 28'$ as in CH_4 , respectively. The synthesis of cubane (C_8H_8), in 1964, and $C_{20}H_{20}$ molecule having a dodecahedron structure, in 1983, started a new era in the chemistry of carbon compounds since in these compounds the carbon bond angles are different from those due to standard hybridizations. Moreover, carbon nanostructures also require different bonding angles.

16.9.1 Synthesis

Carbon clusters are produced by laser evaporation of carbon using laser techniques (Fig. 16.1). When a high intensity laser beam is incident on a carbon substrate in the presence of helium gas, a neutral cluster beam of carbon which when photoionized by UV – laser radiation, carbon clusters are formed. Molecular orbital calculations of the structure of small clusters show the possibility of linear or closed non-planer cyclic geometries as shown in Fig. 16.6. The linear structures form when the number of atoms n is odd and closed structures form when n is even.

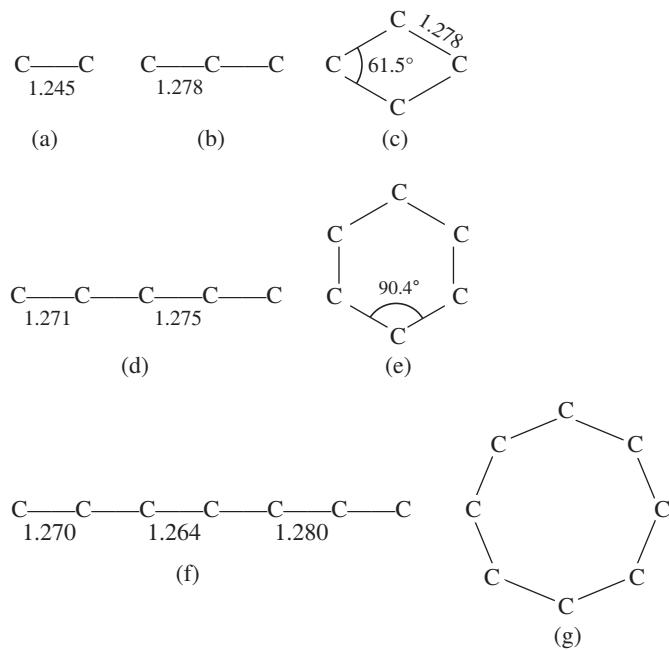


Fig. 16.6 Small carbon clusters.

It has been found from the mass spectrum of carbon clusters that structures with 3, 11, 15, 19, and 23 carbons have the usual bond angles and are more prominent and stable. Again, it is found that the angles in closed structures are different from those of conventional hybridizations. Also, it is noticed that the mass peak at 60 is very intense.

C₆₀ fullerenes

From a study of light transmission through smoke like graphite particles in a helium gas environment, Huffman and Kratschmer obtained conclusive evidence of a new molecule consisting of 60 carbon atoms bonded in the shape of a sphere. The C₆₀ molecule has been named **fullerene** (also called **buckyball**) after the architect Buckminster Fuller who designed the geodesic dome that resembles the structure of C₆₀. It has 12 pentagonal and 20 hexagonal faces symmetrically arranged to form a molecular ball. These ball-like molecules bind together to form a crystal lattice having face centered cubic structure. In the lattice, the nearest neighbour separation between C₆₀ molecules is 1nm. They are held together by the weak Van der Waal's forces. Single crystals of C₆₀ molecules can be grown by slow evaporation from benzene solution.

Larger and smaller fullerenes

The second spherical carbon molecule in the same group is the rugby ball, C₇₀, whose structure has extra six membered carbon rings. Larger fullerenes such as C₇₀, C₇₆, C₈₀, C₈₄ beyond C₁₂₀ have been characterized. A lot more structures are possible with five membered rings in different positions and sometimes together. Smaller fullerenes such as C₂₀, C₂₂, and so on have also been discovered. An interesting aspect of these fullerenes is that they can behave like superconductors at high temperatures.

Endohedral and exahedral fullerenes

Atoms can be placed inside a fullerene ball and such atoms are said to be **endohedral**. Endohedrals with three metal atoms have been trapped in C₆₀, whereas in larger fullerenes the number can be more. Atoms can be bound to fullerenes outside the ball as salts, if the fullerene can gain electrons. In this case, the cation is said to be **exahedral**.

16.10 CARBON NANOTUBES

A more interesting nanostructure with large application potential is the carbon nanotube. A carbon nanotube (CNT) can be thought of as a sheet of graphite rolled in the form of a tube with the ends of the sheet closed by bonds. Figure 16.7(a) shows the structure of a tube formed by rolling the graphite sheet about an axis parallel to the C-C bonds. Figure 16.7 also shows the zigzag and chiral structure.

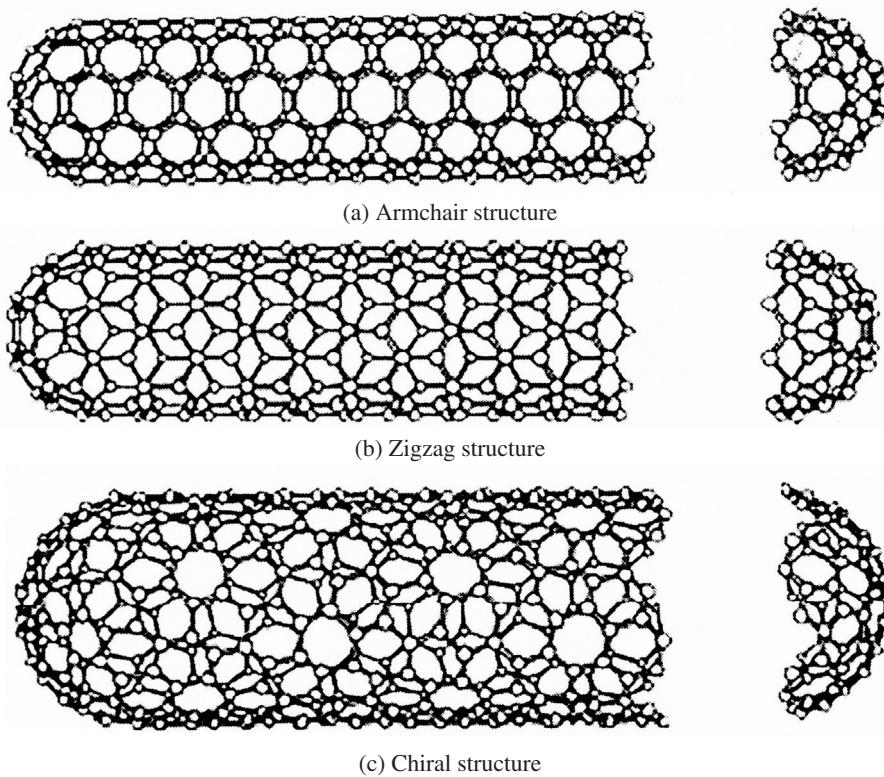


Fig. 16.7 Structure of single walled carbon nanotubes.

16.10.1 Types of Nanotubes

There are two types of nanotubes—single walled nanotube (SWNT) and multi walled nanotube (MWNT). The single walled nanotube resembles a pipe that is capped at each end. The multi walled nano tube resembles many pipes nested within each other. A single walled nanotube can have a diameter of about 2 nm and a length of 100 μm , effectively making it a one dimensional structure called **nanowire**. The properties of both are generally the same. The manufacturing of single walled nanotubes is more difficult than that of multi walled nanotubes.

16.11 FABRICATION OF CARBON NANOTUBES

A number of methods have been developed for the fabrication of carbon nanotubes in sizable quantities such as laser evaporation, carbon arc methods, high pressure carbon monoxide method, and chemical vapour deposition. Some of these processes take place in vacuum or at low pressures with process gases. We shall briefly discuss two of the commonly used procedures in the following sections.

16.11.1 Laser Evaporation Method

Figure 16.8 illustrates the schematic arrangement of the set up used for making carbon nanotubes by laser evaporation. A quartz tube containing flowing argon gas and a graphite target with arrangement for irradiation by pulsed laser are kept inside a furnace. In the other end of the quartz tube outside the furnace, a water cooled copper collector is provided. The graphite target is heated to about 1200°C . The graphite target contains a small amount of cobalt and nickel acting as catalytic nucleation sites for the formation of the tubes. When the intense pulsed laser beam is incident on the target, carbon evaporates from the graphite. The argon then sweeps the carbon atoms from the high temperature zone to the colder copper collector where they condense into nanotubes. Tubes of 10–20 nm in diameter and 100 μm long can be fabricated by this arrangement.

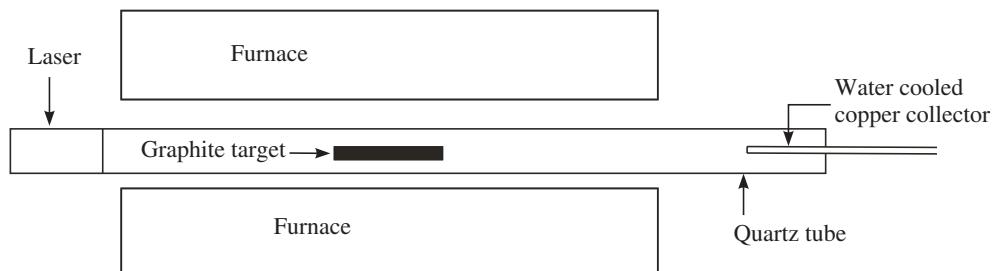


Fig. 16.8 Fabrication of carbon nanotubes by laser evaporation method.

Arrangement is also provided for heat treatment in vacuum, at 1000°C to remove the C_{60} and other fullerenes. By varying the growth temperature, the catalyst composition, and other process parameters, the average nanotube diameter and size can be varied.

16.11.2 Chemical Vapour Deposition

This method is also referred to as vapour-liquid-solid (VLS) growth. The VLS apparatus is also a tube furnace with arrangement for the flow of gases while it is being heated to high temperatures. This procedure involves the decomposition of a hydrocarbon such as methane (CH_4) at 1100°C . The gas decomposes, forming carbon atoms which condense on a cooler substrate containing a catalyst such as iron, cobalt, or nickel. This method gives nanotubes with open ends, which is not common by other methods. The single walled nanotubes grow at the sites of the metal catalyst. The catalyst particles generally stay at the tip of the growing single walled nanotubes during the growth process. Methane is selected because of its high thermal stability and its ability to retard the formation of amorphous carbon in the tube.

16.12 NANOWIRES AND NANORIBBONS

Nanowires are solid cylindrical material (not hollow as tubes) with a uniform diameter of a few tens of nanometers or less. They are usually synthesized from standard semiconductors Si, Ge, GaAs, GaP, GaN, InAs, InP, ZnS, ZnSe, CdS, CdSe, and mixed compounds. A nanowire with

segments of different materials along their length is called a **barcoded nanowire**. These can be synthesized with segments differing in electronic properties (conductors, semiconductors) or optically distinguishable patterns, with each segment as small as tens of nanometers.

Nanoribbons have a uniform rectangular cross section with well defined crystal structures and synthesized from the oxides of metals and semiconductors. Nanoribbons are perfectly single crystalline and essentially free from dislocations. Hence, they provide an ideal standard for the systematic study of electrical, thermal, optical, and transport processes in semiconductor nanomaterials.

16.13 PROPERTIES OF NANOMATERIALS

Nanomaterials have exceptional electrical, mechanical, thermal, and optical properties. We shall discuss some of them here.

16.13.1 Metal Nanoclusters

Geometric structure

The crystal structure of large nanoparticles is the same as bulk structure with slightly different lattice parameters. However, in the case of clusters with a small number of nanoparticles, the structure is often different from the bulk one. In the case of 3-5 nm gold particles, it is found that they have an icosahedral structure rather than the bulk fcc structure. The possibility of a structural change is noticed even in an aluminium cluster of 13 atoms. Indium clusters undergo a change of structure when the size is smaller than 5.5 nm.

Electronic structure

In a metal, the conduction band is not filled completely. The number of energy levels in a given interval of energy is the density of states. When a metal having bulk properties is reduced in size to a few hundred atoms, the density of states in the conduction band changes drastically. The continuous density of states is replaced by a set of discrete energy levels. The small cluster is analogous to a molecule having discrete energy levels with bonding and antibonding orbitals.

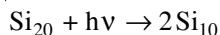
16.13.2 Nanoparticles of Semiconducting Materials

Optical properties

Nanoparticles made of germanium, silicon, or cadmium show peculiar optical properties, but they are not semiconductors. There is a significant shift in the optical absorption spectra towards shorter wavelength as the particle size is reduced. In a bulk semiconductor, a bound electron-hole pair, called an **exciton**, is produced by a photon having an energy greater than that of the band gap of the material. The exciton moves through the lattice as a single entity. When the size of the nanoparticles is larger than the radius of the electron-hole pair, there is a blue shift of the absorption spectrum. However, when the size of the nanoparticle is smaller, the motion of electron and hole become independent and the exciton does not exist. The hole and the electron have their own sets of energy levels.

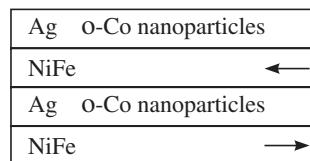
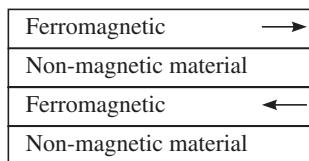
Photofragmentation

It is observed that nanoparticles of germanium and silicon undergo fragmentation when subjected to laser light from a Q-switched Nd:YAG laser. The products depend on the size of the cluster, intensity, and the wavelength of light. Some of the fissions observed are



16.13.3 Giant Magnetoresistance

The application of a DC magnetic field changes the resistance of a material. This phenomenon is called **magnetoresistance**. Magnetoresistance occurs in metals at very high magnetic field and low temperatures. Materials fabricated by depositing on a substrate alternate layers of nanometer thickness of a ferromagnetic material and a nonferromagnetic metal are called **giant magnetoresistance** materials. Figure 16.9(a) shows a schematic representation of the layered structure and the alternating orientation of the magnetization of the ferromagnetic layer. Films made of alternating layers of iron and chromium, cobalt and copper are some of the familiar examples. In the case of a carbon nanotube, because the resistance decreases with the increase of magnetic field, it is a case of negative magnetoresistance effect. Hybrid systems consisting of nanoparticles in metal matrices sandwiched between metal magnetic layers also exhibit similar behaviour [as shown in Fig. 16.9(b)]. Materials made of single-domain ferromagnetic nanoparticles with randomly oriented magnetizations embedded in a non-magnetic matrix also display giant magnetoresistance.



(a) Schematic representation of the layered structure and the alternating orientation of the magnetization in the ferromagnetic layer.

(b) Hybrid system consisting of cobalt nanoparticles in a silver matrix sandwiched between nickel-iron magnetic layers with alternating magnetizations (arrows).

Fig. 16.9 Magnetoresistance.

16.13.4 Carbon Nanotubes

Electrical conductivity

Carbon nanotubes are metallic or semiconducting depending on the diameter and chirality of the tube. Synthesis gives a mixture of tubes, two-thirds of which are semi conducting and one-third metallic. The metallic tubes have the armchair structure. A plot of the energy gap of the semiconducting chiral carbon nanotubes versus the reciprocal of the diameter is shown in Fig. 16.10. From the figure it is obvious that the band gap decreases as the diameter of the tube

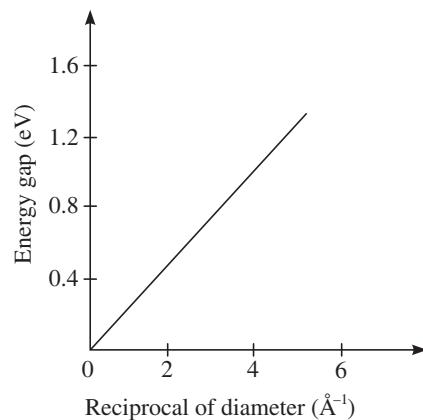


Fig. 16.10 The energy band gap of semiconducting chiral carbon nanotube versus the reciprocal of the diameter of the tube.

increases. In the metallic state, the conductivity of the nanotubes is very high. It is estimated that they can carry a billion amps/cm². Copper wires fail at one million amps/cm² because of resistive heating, which melts the wire.

Thermal conductivity

Nanotubes are also good conductors of heat. They have a very high thermal conductivity, almost twice that of a diamond.

Mechanical properties

Carbon nanotubes are very strong. If a weight is attached to the end of a nanowire nailed to the roof of a room, the wire stretches. Carbon nanotubes have Young's moduli ranging from 1.28 to 1.8 TPa (terapascal). Young's modulus of steel is 0.21 TPa. This implies that the carbon nanotube has a Young's modulus that is eight times the Young's modulus of steel. Multi walled nanotubes of 200 nm diameter have modulus of 0.6 TPa. The tensile strength of single walled nanotubes is about 20 times that of steel.

Field emission of electrons

Carbon nanotubes can have length to diameter ratio as large as 10^5 . Hence, field emission of electrons can be induced from the tip of long metallic carbon nanotubes in the presence of moderate electric fields. Better field emission is possible by introducing adsorbates that lower the ionization potential. Experiments in the presence of CO, CO₂, H₂, and water vapour showed that adsorbed water can significantly enhance the emission current.

16.14 APPLICATIONS

Nanomaterials have very wide ranging applications. We discuss briefly some of them in this section:

Nanotechnology and the Energy Challenge

One of the most important technical challenges faced by the world today is sustainable energy supply. Nanotechnology can play a key role in the development of sturdier fuel cells, improved membrane technology, and storing hydrogen in fuel cells. Nanostructured ceramic membranes derived from metal oxane nanoparticles can improve the efficiency of fuel cells. Titanium dioxide nanotube arrays have already been developed for generating hydrogen by splitting water using sunlight. With the increase in the number of times the process is repeated the potential of the process will increase as a clean energy resource. The hydrogen thus generated can also be stored in nanoporous titanium and carbon nanotube assemblies. These nanomaterials are powerful enough to maintain hydrogen for use in vehicles.

Solar energy can be generated through photovoltaic semiconductor junctions, plants, and through splitting of water using sunlight. The advanced photovoltaic solar cells absorb a photon of sunlight, and the energy gets converted to one electron and the rest is lost as heat. It has already been found that quantum dots can produce as many as three electrons from one high energy photon of sunlight. This nanotechnology method can boost the efficiency of solar cells from the present 20–25 per cent to 60–65 per cent.

Machine tools

In general, nanomaterials are harder than conventional materials. Therefore, cutting tools and drills made of nanocrystalline materials such as titanium carbide, tungsten carbide, and tantalum carbide are harder and wear resistant. Nanocrystalline silicon nitride and silicon carbide are extremely useful for the manufacture of high strength springs, ball bearings, etc. They are also used as components in high temperature furnaces.

Batteries

Nanocrystalline materials synthesized by sol-gel method result in aerogels. Because of their aerogel structure they are used as separator plates in new generation batteries. These materials can hold considerably more energy as compared to the conventional plates. Frequent recharging is not needed with nickel metal hydride (Ni-MH) batteries made of nanocrystalline nickel and metal hydrides. The lifetime of such batteries is also much longer.

High power magnets

The increased surface area per unit volume available with nanocrystalline materials increases the strength of a magnet. Magnets made of nanocrystalline yttrium-samarium-cobalt have unusual magnetic properties. These high power rare-earth magnets are extremely useful for applications such as motors for ships, ultra-sensitive analytical instruments, magnetic resonance imaging, power generators, automobile alternators, etc.

Insulation materials

Nanocrystalline materials synthesized by the sol-gel process results in aerogels. Aerogels are three dimensional continuous network of particles with air trapped at their interstices. They are porous and extremely light. They are used for insulation in offices and homes, thereby reducing electricity charges and environmental pollution.

Shielding material for electromagnetic radiation

As carbon nanotubes have very high electrical conductivity they are poor transmitters of electromagnetic energy. A plastic composite of carbon nanotubes can provide light weight shielding material for electromagnetic radiation. This would be extremely useful, especially for military applications.

Switching component in computers

FET, the switching component of computers, based on semiconducting carbon nanotubes connecting two gold electrodes has been demonstrated. It is estimated that the switching time of these devices will be 10^4 times faster than present processors. The smaller size of the device with nanotube allows more switches to be packed on a chip. Another idea is to make a computer out of carbon nanotubes. The computer would be an array of parallel nanotubes on a substrate. Above this, another array of carbon nanotubes would be oriented perpendicular to the tubes, but not touching the lower array. Each tube would be connected to a metal electrode. It is estimated that 10^{12} switches could fit on a cm^2 chip, which is much higher than that in the present ones.

Chemical sensors

SnO_2 nanoribbon sensors are highly effective in detecting even very small amounts of harmful gases such as NO_2 . On absorption of these gases, the electrical conductance of the sample decreases. Electron withdrawing groups such as NO_2 and O_2 are expected to deplete the conduction electron population in the nanoribbon, thereby decreasing the electrical conductance. An interesting aspect regarding the sensors is that it is possible to remove the adsorbates by shining ultraviolet light on it. This sensing device has advantages over the conventional thin film oxide sensors: low operating temperatures, high active surface to volume ratio, and no ill defined coarse grain boundaries.

Nanoelectrodes

Nanoscale electrodes can display altered properties. Carbon nanotubes and metallic arrays have been used as electrodes to promote electron transfer in redox reactions. The increased available reactive surface area and focussing of the electric field due to the rod like geometrical configuration are responsible for the better performance.

Biomedical implants

Biomedical implants are very common and are used for orthopedic fixation, dental purposes, vascular and cardiovascular organ purposes, etc. Nanotechnology is a technique to work at the molecular level atom by atom to create large structures having new molecular organization with at least one dimension measured in nanometers. Surface properties such as area, charge, and morphology depend on the grain size or fibre size of the material. Cells in our body are certainly accustomed to interacting with nanostructured surfaces. Because of their smaller size, the metallic particulates require less severe sintering conditions such as temperature and pressure to become fully consolidated. As already mentioned in earlier sections, nanomaterials have the properties required by a biomedical implant such as strength, elastic modulus, ductility,

electrical conductivity, and increased surface roughness. Owing to their improved material properties, the day is not far off when the nanoparticles will replace the conventional ones as a biomedical implant. Studies have already shown the potential for the use of carbon nanofibres as neural prostheses. Recent studies on conventional and nanoparticle titania / PLGA (poly lactic-co-glycolic acid) composites for use as bone prostheses demonstrated much higher adhesion of both bone producing cells and cartilage synthesizing cells by the nano cells.

16.15 MEDICAL APPLICATIONS

The gel electrophoresis method was generally used to separate fragments of DNA for analysis. Silicon based nanostructures with pores comparable to the size of a large DNA molecule can also be used for the study of other large organic molecules, including proteins, carbohydrates, and lipids. For these, the electrophoresis approach is not an efficient one.

Another medical application is the first response **home pregnancy test** which uses micrometer sized latex particles in conjunction with gold nanoparticles of less than 50 nm diameter. When the urine of a pregnant woman is mixed with the micro and nanoparticles combination, coagulation takes place due to the presence of a hormone, resulting in a pink colour.

Carbon nanotubes have shown very good results as a biochemically compatible nanoelectronic interface to biomolecules. Carbon nanotubes with an immobilized protein at one end offer a number of advantages. It maintains the protein's native conformation to the maximum possible extent. Its high electrical conductivity provides an efficient conduit for electronic signals from or to the protein. Its sharp curvature at the tip greatly increases the electrical field, and thereby increases the electron transfer between the protein and the nanotube exponentially. Activated carboxyl groups are readily produced via acid treatment at the open end of the carbon nanotube. This facilitates its conjugation with biomolecules localized at the tip.

Some other applications of nanomaterials are as follows:

1. Nanogold particles are used extensively as staining agents in biological electron microscopy.
2. Iron oxide nanoparticles are used as contrast agents in MRI scan. They are much more potent than other magnetic contrast agents.
3. To prevent wastage of thermal energy, the engine cylinders of motor cars are coated with nanocrystalline ceramics such as zirconia and alumina.
4. Nanocrystalline zinc-selenide, zinc sulphide, cadmium sulphide and lead telluride are better phosphors than the usual ones.
5. Nanostructures are not continuous, and therefore, they are used in place of venetian blinds.
6. Nano-thick coatings of clear materials such as nano-rust and nano-sand (iron and silicon oxide layers) give colour due to the wave from the front and back surfaces, producing interference. Different thickness gives different colours. These paints are in use in cars, jewellery and other decorations.

REVIEW QUESTIONS

- 16.1** What are nanomaterials? Why do these materials exhibit properties different from those of their classical counterparts?
- 16.2** Outline briefly the top down and bottom up methods used for the synthesis of nanoparticles.
- 16.3** What is a carbon nanotube? Explain the different types of carbon nanotubes.
- 16.4** Distinguish between a nanoribbon and a nanowire.
- 16.5** Outline some of the properties of carbon nanotubes.
- 16.6** Explain some of the properties of nanomaterials.
- 16.7** What is a bucky ball?
- 16.8** Nanotechnology can boost the efficiency of solar cells from the present 20–25% to 60–65%. Comment.
- 16.9** Explain the use of nanomaterials as an energy storage device.
- 16.10** Comment on the statement “Owing to their improved material properties, the day is not far off for the nanoparticles to replace the conventional ones as biomedical implants”.
- 16.11** Explain the bottom up method for the preparation of nanomaterials. Outline with a diagram the synthesis of a nanomaterial by this procedure. Explain some of its properties.
- 16.12** Describe with the principle the chemical vapour deposition method for the preparation of nanomaterials. How is this method used to grow structures vertically? Mention the applications of nanomaterials.
- 16.13** What is a carbon nanotube? How are carbon nanotubes fabricated? Explain some of its properties and applications.
- 16.14** Outline some of the properties of nanomaterials and explain the applications based on these properties.

C H A P T E R

17

Ultrasonics

Ultrasonics whose frequencies are outside the audible range have a number of applications in non-destructive testing, in medicine, and in other fields. In this chapter, we consider certain aspects of its production, properties, and applications.

17.1 ULTRASONIC WAVES

Human ears are sensitive to sounds with a frequency upto 20 kHz. Sound waves of frequency greater than 20 kHz are called **ultrasonics**. The upper frequency limit is upto the range of 500 MHz. Many animals can hear ultrasonic frequencies, for example, dogs can hear sounds as high as 50 kHz, and bats can detect frequencies as high as 100 kHz. Ultrasonic waves have number of applications in engineering, medicine and other fields. As the speed of sound is about 350 ms^{-1} at room temperature, ultrasonic waves having a frequency of 5 MHz have wavelength of $7 \times 10^{-5} \text{ m}$. Because of its high frequency, the penetrating power is also high. Sound waves whose frequencies are below the audible range (less than 20 Hz) are called **infrasonics**. Sources of infrasonic waves include earthquakes, volcanoes, thunder and waves produced by vibrating heavy machines.

17.2 PRODUCTION OF ULTRASONICS—MAGNETOSTRICTION METHOD

17.2.1 Principle

When a ferromagnetic rod is subjected to a magnetic field parallel to its length, the length increases or decreases. This effect is called **magnetostriiction**. If the length increases, it is called

positive magnetostriiction and if it decreases, it is called **negative magnetostriiction**. The effect is strong on metals like iron, cobalt, nickel and their alloys. The change in length depends on the intensity of the magnetic field applied. However, it is independent of the direction of the field. Hence, when an alternating field is applied, the rod suffers the change in length during each half cycle of the alternating current (a.c.). Consequently, the rod vibrates with a frequency twice that of the frequency of a.c. If the natural frequency of the rod coincides with the frequency of the a.c., due to resonance the rod vibrates with considerable amplitude. If the frequency of the a.c. is high, ultrasonic waves are produced from the ends of the rod.

We shall next find an expression for the resonant frequency f for the longitudinal vibrations of a bar of length l . The velocity of ultrasonic waves in a bar is given by

$$v = \sqrt{\frac{Y}{\rho}} \quad (17.1)$$

where Y is the Young's modulus and ρ is the density of the material. If the rod vibrates in its fundamental, the length $l = \lambda/2$, and therefore,

$$\begin{aligned} v &= f\lambda \quad \text{or} \quad \sqrt{\frac{Y}{\rho}} = f2l \\ f &= \frac{1}{2l} \sqrt{\frac{Y}{\rho}} \end{aligned} \quad (17.2)$$

17.2.2 Experimental Arrangement

In general, the alternating magnetic field is generated by an electronic oscillator and its frequency is controlled by a variable capacitor. The generally used oscillator is the Hartley oscillator, its circuit diagram is given in Fig. 17.1. Frequencies ranging from a few hundred hertz to about 300 kHz can be produced. Inductors L_1 and L_2 and the variable capacitor C together form the tank circuit. Invariably it uses a single tapped coil having two parts L_1 and L_2 instead of two separate coils. A rod of ferromagnetic material is placed inside the coil as in figure and clamped at the middle. When the oscillator is switched on, capacitor C gets charged. When the capacitor is fully charged it discharges through L_1 and L_2 setting up oscillations. In the high frequency current, the rod gets magnetized and then demagnetized and so on. Consequently, the length of the rod changes and its free ends produce high frequency vibrations of ultrasonics. When the frequency of oscillation produced by the oscillator equals the natural frequency of the rod, resonance occurs and the amplitude of vibrations becomes very large.

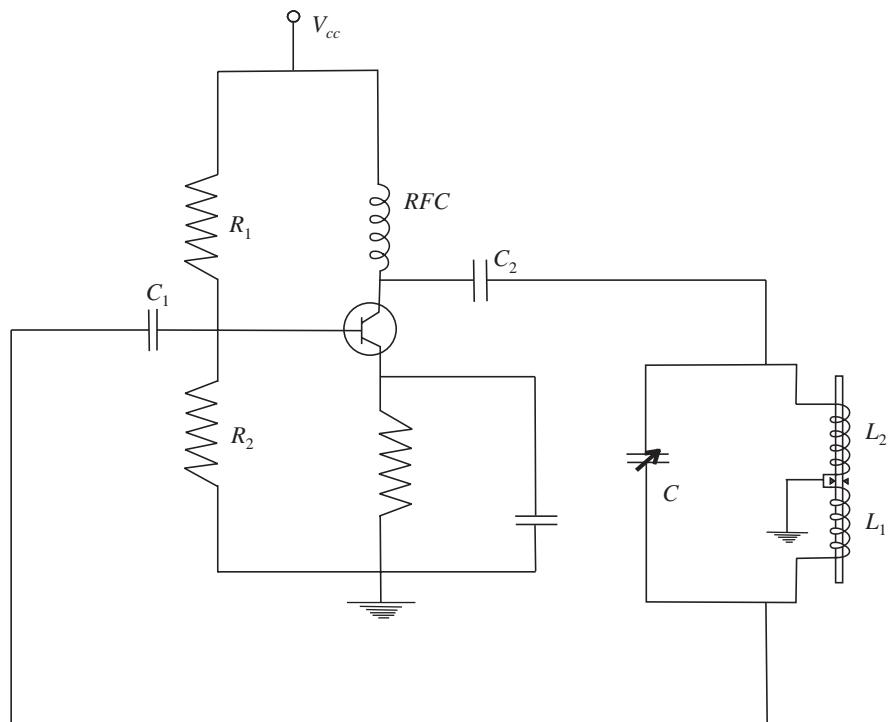


Fig. 17.1 Production of ultrasonics - Magnetostriction method.

17.3 PRODUCTION OF ULTRASONICS – PIEZOELECTRIC METHOD

17.3.1 Principle

In certain crystals, the application of an external mechanical stress between two opposite faces induces a net dipole moment which produces polarization charges on the surfaces of the crystal. In other words, application of stress gives rise to a potential difference. This phenomenon is called **piezoelectric effect**. Conversely, if an alternating voltage is applied between two opposite faces of a crystal like quartz or rochelle salt or barium titanate, corresponding changes in the dimension of the other pair of faces occur, thereby generating elastic waves. Thus, the inverse piezoelectric effect will give rise to ultrasonic waves. An important point to be noted is that the crystal should not have an inversion centre.

For applying the alternate voltage, a Hartley oscillator is used. A thin slice of the crystal cut with its opposite faces perpendicular to the optic axis is placed between two metal plates which are connected to the output of the oscillator as shown in Fig. 17.2. When the oscillator oscillates, high frequency alternating voltages are impressed on the plates A and B . Inverse piezoelectric effect takes place and the crystal vibrates, thereby generating ultrasonic wave in the nearby medium. The natural frequency of the crystal f is the same as Eq. (17.2). When the frequency of the a.c. matches the frequency of the crystal, resonant vibrations are produced with very high amplitudes. Ultrasonic frequencies as high as 500 MHz can be obtained by this method.

17.3.2 Experimental Arrangement

The circuit used for production is shown in Fig. 17.2. The coil L_1 and L_2 are connected in series. The coil L_3 is inductively coupled to $L_1 - L_2$ combination. The crystal P is placed between two metal plates A and B . When V_{cc} is switched on, collector current starts rising and charges the capacitor C . When the capacitor is fully charged, it discharges through the coils L_1 and L_2 setting up oscillations in the tank circuit. Consequently, high frequency alternating voltages are set up in the plates A and B . As a result the crystal contracts and expands periodically, thereby producing compressions and rarefactions in the surrounding medium. When the frequency of oscillation produced by the oscillator is equal to the natural frequency of the crystal, resonance occurs producing ultrasonics of very large amplitude.

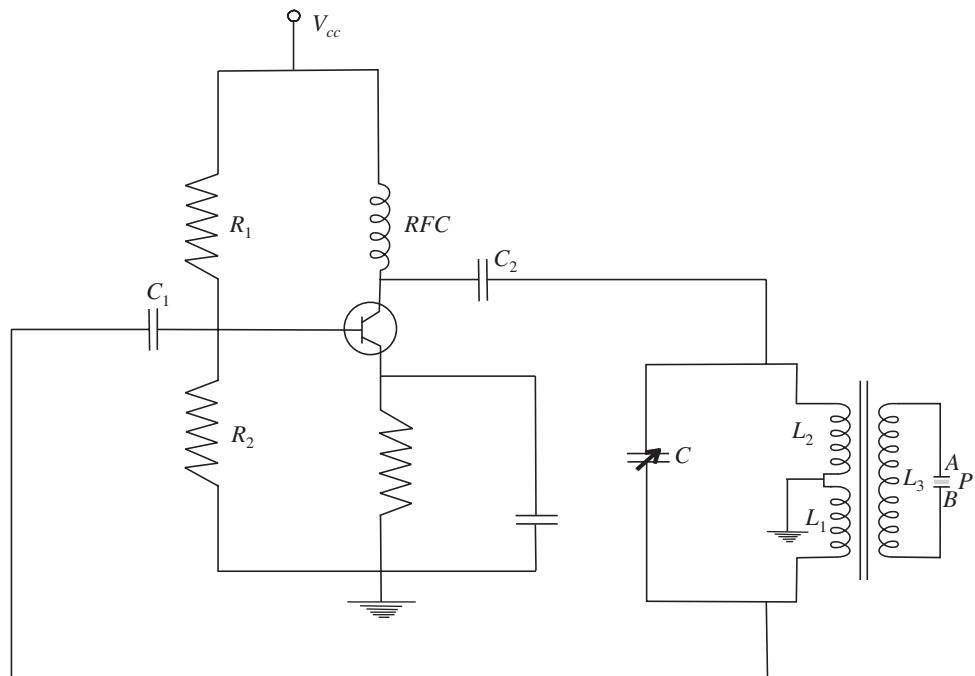


Fig. 17.2 Production of ultrasonics by piezoelectric method.

17.4 DETECTION OF ULTRASONIC WAVES

Thermal detectors

In this method, a probe of platinum wire is moved in the region to be tested for ultrasonic waves. The temperature of the medium through which the ultrasonic wave travels changes due to alternate compressions and rarefactions. At antinodes the temperature remains constant, whereas at the nodes the temperature changes. Consequently, the resistance of the wire changes at the nodes and remains constant at antinodes. The change in the resistance of the wire can be detected using sensitive bridge arrangement.

Piezoelectric detectors

Quartz used for the production of ultrasonics can be used for its detection also. When one pair of faces of a quartz crystal is subjected to ultrasonic waves, varying electric charges are produced on the other pair of faces. These charges are amplified and then detected using electronic circuits.

Kundt's tube method

Kundt's tube can be used to detect ultrasonic waves as done for audible sound waves. Kundt's tube is a long cylindrical glass tube 3 to 4 cm in diameter and open at both ends. One end is loosely closed with a cork and the other end has a provision to fix up a movable piston. When ultrasonic waves are passed through the tube, the lycopodium powder inside the tube collects in the form of heaps at nodes and is blown off at the antinodes.

Sensitive flame method

A sensitive narrow flame is moved in the region to be tested for ultrasonic waves. The flame remains stationary at antinodes and flickers at node.

17.5 PROPERTIES OF ULTRASONICS

- (i) The speed of sound waves in air $\cong 350 \text{ ms}^{-1}$. For a wave having a frequency of about 10 MHz, the wavelength is given by

$$\lambda = \frac{v}{f} = \frac{350 \text{ ms}^{-1}}{10 \times 10^6 \text{ s}^{-1}} = 3.5 \times 10^{-5} \text{ m} \quad (17.3)$$

a very small value. Since the wavelength is very small diffraction effects will be negligible. Hence, ultrasonic waves can be transmitted over long distances without much loss of energy. This property is used extensively for number of ultrasonic guided techniques.

- (ii) The energy per unit volume of ultrasonic waves is directly proportional to the square of its frequency. Hence, they are highly energetic. The power is of the order of watts and in certain cases several kilowatts. Applications based on high intensity of ultrasonics are enormous.
- (iii) When passed through liquids and gases they produce a disruptive effect which has lot of practical applications.
- (iv) The waves produce emulsion at the interface between two liquids, and hence, can be used for mixing oil and water. This property is extensively used in the preparation of food products, paints, cosmetics and so on.
- (v) When passed through substances, ultrasonic waves produce cavitation. This effect is utilized in cleaning, degreasing and soldering of aluminium and its alloys.
- (vi) Ultrasonic waves of high intensity produce chemical changes when passed through certain liquids. Certain chemical reactions are initiated by ultrasonic waves. In certain others it accelerates the chemical reaction.

17.6 ACOUSTIC GRATING

When ultrasonic waves are passed through a liquid, due to compressions and rarefactions, the density varies from layer to layer. The liquid density and consequently the refractive index is maximum at nodes and minimum at antinodes. As a result, the regions of greater and lesser refractive indices act like an **acoustic grating**. Under this condition, if monochromatic light is passed through the liquid at right angles to the direction of propagation of ultrasonic waves, the liquid behaves like a grating, resulting in the diffraction of light. If d is the distance between two adjacent nodal or antinodal planes, we have

$$d \sin \theta = n\lambda \quad (17.4)$$

where θ is the angle of diffraction, λ is the wavelength of the monochromatic light and n is the order of diffraction. Since d is the distance between adjacent nodes or antinodes

$$2d = \lambda_u \quad \text{or} \quad d = \frac{\lambda_u}{2} \quad (17.5)$$

where λ_u is the wavelength of ultrasonic waves. With this value of d , Eq. (17.5) gives

$$\begin{aligned} \frac{\lambda_u}{2} \sin \theta &= n\lambda \\ \lambda_u &= \frac{2n\lambda}{\sin \theta} \end{aligned}$$

Using the frequency of the ultrasonic wave, one can evaluate the velocity of the ultrasonic waves in the liquid. And using the velocity, one can evaluate the compressibility of the liquid.

17.7 NON-DESTRUCTIVE TESTING

When no permanent changes take place in the properties of materials during the process of testing, it is called Non-Destructive Testing (NDT). Low intensity ultrasonics is generally used for NDT. Number of methods such as pulse echo method, progressive continuous wave method, stationary or resonance method, reverberation method are available. All these methods are based on the measurements of acoustic velocities and attenuation. The involved frequency range is roughly 0.5 MHz to 10 MHz. We shall discuss here the pulse echo and transmission methods.

17.7.1 Pulse Echo Method–Flaw Detection

A schematic representation of the method is illustrated in Fig. 17.3. The transducer (A) is excited by an electric pulse from the generator. The transducer emits a short pulse of ultrasonic waves in the specimen under test. At the same instant, the ultrasonic pulse passes directly to the CRO through an amplifier and a large peak P is observed on the screen. The ultrasonic pulse reflected from the opposite boundary of the sample is received by the transducer, amplified and fed to the CRO which appears as peak Q in the screen. When the pulse pass through the sample, if there is a flaw, partial reflection takes place at that discontinuity which produces an additional peak in the CRO. From the position of the additional peak in the CRO, the exact location of flaw can be estimated. The major advantages of the method are as follows:

- (i) Access to one side of the test specimen is sufficient.
- (ii) Distance of the defect from the transducer can be measured.
- (iii) The shape of the specimen is not important.
- (iv) Has only one coupling point. Hence less expensive and less chances of error.

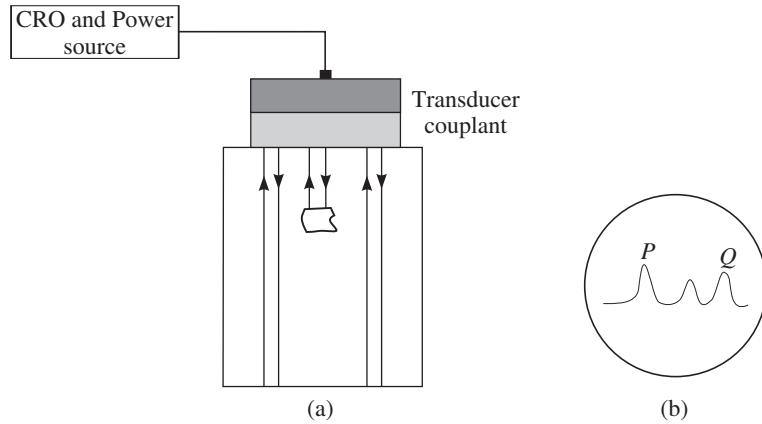


Fig. 17.3 Ultrasonic flaw detector.

17.7.2 Transmission system

An alternate procedure is the transmission system in which the ultrasonic pulse is transmitted through the system as shown in Fig. 17.4. In this method the transmission probe is placed in contact with the test specimen using a liquid coupler and the receiving probe is placed on the opposite side of the specimen as shown in the figure. If a defect is present in the test specimen

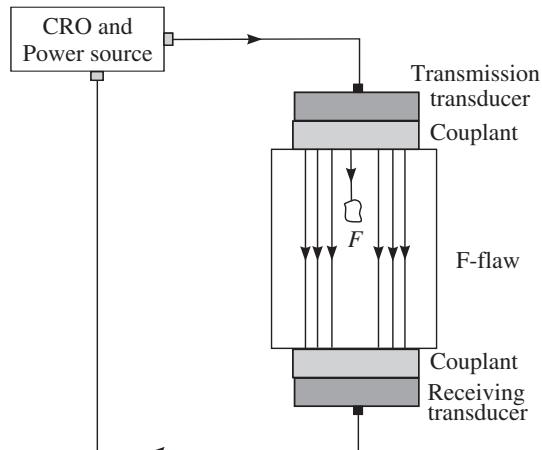


Fig. 17.4 Transmission ultrasonic system for flaw detection.

partial reflection at the flaw causes a reduction in the strength of the signal received. Some of the disadvantages are as follows:

- (i) The specimen must have parallel sides and one has to approach the specimen from both the sides
- (ii) Two transducers are needed
- (iii) The two transducers must be exactly opposite
- (iv) Only the presence of flaw can be studied, not the exact location

17.7.3 Thickness Gauging

For ultrasonic thickness measurement either pulse echo or resonance method is used. In the pulse echo procedure, the principle is similar to flaw detection. First the gauge is calibrated with a standard sample. The time taken for travelling the distance (thickness) is measured by electronic circuits and the thickness is estimated. The important advantage is that access to one surface is sufficient for the thickness monitoring process. The procedure is used for various processes like measuring the thickness of a ship's hull for external corrosion from inside the ship, for the assessment of internal corrosion of pipes, for thickness measurement of fat on human body and so on.

17.7.4 Sonar

Sonar stands for **s**ound **n**avigation for **r**anging. Submerged objects such as submarines, rocks, icebergs, shoal of fish, etc., can be detected. A sharp ultrasonics beam is directed in various directions into the sea. By noting the time taken by the reflected beam an idea of the distance of the reflecting body is obtained. The change in frequency of the echo signal due to doppler effect gives the velocity of the body. The depth of the sea can also be estimated from the time interval between the sending of the wave and the reception of the reflected wave.

17.8 MEDICAL APPLICATIONS OF ULTRASONICS

17.8.1 Ultrasound Imaging

Ultrasonics have number of applications in medicine. It is used both in the process of diagnosis and in various treatments. The diagnostic use involves the non-invasive imaging of the internal organs or the structures of the body. Ultrasound imaging systems generally use the pulse echo and pulsed doppler techniques. In the **pulse echo technique** an ultrasound is directed into the body and its reflection from organs and other structures and lesions in the body are detected. In **pulsed doppler**, a high frequency sound beam is send and the echo is received. When the object that reflects the sound is in motion, there will be a change in the frequency of the reflected signal. From the change, one can estimate the direction and blood velocity. With the help of this one can find the deposition of lipid and cholesterol particles in the walls of artery, malignancy, vessel narrowing, etc. Certain kinds of tissue or fluid are not detected in x-ray photographs, but ultrasound waves are reflected from their boundaries.

17.8.2 Echocardiography

Echocardiography is another area where ultrasonics technique is used which helps to image cardiac structures dynamically. A piezoelectric transducer is placed on the outer chest and an ultrasonic beam is directed towards the heart. The echoes from the heart muscle are collected by the same transducer. A single transducer thus, acts as a transmitter and receiver, alternately. By changing the position of the transducer we can get reflections from the desired areas on the heart. An aqueous gel is used to couple the transducer to the skin. The beam from the transducer can reach upto a depth of 5 – 10 cm. The amplified signals are displayed in an oscilloscope.

In recent years number of technological developments have taken place in the field of echocardiography. The two dimensional echocardiography has permitted the identification of holes, the study of volume of heart and valvular lesions. The continuous wave doppler arrangement allows the measurement of pressure gradient along the line of probe. A transducer placed within the esophagus by a flexible gastroscope increases the use of echocardiographic examinations in the operating room.

17.8.3 Ultrasound Guided Technique

Another field in which ultrasonics technique is applied is the ultrasound guided techniques which makes possible the execution of various invasive procedures like cyst puncture, obstetric puncture, treatment of the lymph nodes, prostate, etc. Tapping of fluid from the abdomen, fluid drainage and abscess drainage are some of the common procedures. The process also helps surgeons by localising masses, foreign bodies and stones within organs during surgery.

17.9 OTHER APPLICATIONS

Cavitation

When ultrasonic waves of high intensity travels through a liquid, shock waves are generated. Consequently, the local temperature rises to extremely high values for very short time. This is called **cavitation**. Cavitation process is used in emulsification of two immixable liquids. This process is in use in food industry mainly for the preparation of dairy products, sauces, gravies, synthetic creams and so on. It is also used in the extraction of hops in brewing industry. Tendering of meat, sterilization of milk are some of the other applications.

Effects on suspension

Ultrasonic waves when pass through fluids cause agitation. This effect finds application in the dispersion of fog and smokes, degassing of liquid foods, etc.

In metallurgy

Molten metals which are in the process of cooling are agitated by ultrasonics to refine the grain size, to prevent the formation of cores and to release trapped gases. To get alloys of uniform composition, at the manufacturing stage, it is agitated using a beam of ultrasonics.

Ultrasonic cleaning

In ultrasonic cleaning both cavitation and agitation are involved. Cleaning of large components is carried at low frequencies (18 kHz to 40 kHz), where cavitation is most active. For delicate items such as printed circuit boards, jewellery etc higher frequency range (100 kHz to 1 MHz) is used, for the removal of grease, oil, dust, buffering compounds, etc. At high frequencies agitation is the most active one.

Metal cutting, drilling and welding

Ultrasonic drill with high frequency waves is widely used for drilling in materials like ceramics, glass, germanium, etc. High frequency waves are used for welding and soldering. Ultrasonic welding can be done at room temperature as well. To solder aluminium, ultrasonic waves alongwith electrical soldering iron is a must.

SOLVED EXAMPLES

Example 17.1 The fundamental vibrational frequency of a quartz crystal is 3 MHz. If it is vibrating at resonance, what is its thickness. Young's modulus of quartz = 7.9×10^{10} N/m² and density of quartz = 2650 kg/m³.

Solution: For the resonant frequency, we have

$$\begin{aligned} f &= \frac{1}{2l} \sqrt{\frac{Y}{\rho}} \quad \text{or} \quad l = \frac{1}{2f} \sqrt{\frac{Y}{\rho}} \\ l &= \frac{1}{2(3 \times 10^6 \text{ s}^{-1})} \sqrt{\frac{7.9 \times 10^{10} \text{ N/m}^3}{2650 \text{ kg/m}^3}} \\ &= 0.09 \times 10^{-2} \text{ m} = 0.9 \text{ mm} \end{aligned}$$

REVIEW QUESTIONS

- 17.1** List some of the properties of ultrasonics.
- 17.2** What is magnetostriction?
- 17.3** What is piezoelectric effect?
- 17.4** Explain two important applications of ultrasonics as a tool for non-destructive testing.
- 17.5** Explain how piezoelectric effect is utilized for the production of ultrasonic waves. Explain some of the applications of ultrasonics.
- 17.6** What is magnetostriction? Explain how ultrasonic waves can be produced by using magnetostriction effect. Also explain any two uses of ultrasonics as a non-destructive testing tool.

PROBLEMS

1. A certain piezoelectric crystal of thickness 3 mm produces ultrasonic waves of frequency 400 kHz. Calculate the thickness of the crystal to produce ultrasonic waves of frequency 600 kHz.
2. For a quartz crystal of length 0.04 m, calculate the fundamental frequency of vibration in a piezoelectric oscillator, given that the velocity of longitudinal waves in the crystal is 5.5×10^3 ms⁻¹.

C H A P T E R

18

Acoustics of Buildings

Acoustics is a branch of physics that deals with the process of generation, transmission and reception of sound in a room or in a hall. The area of architectural acoustics dealing with the design and construction of buildings and halls is an important factor in the phenomena. In this chapter, we would discuss certain aspects of architectural acoustics so that the listeners can experience best sound.

18.1 ACOUSTICS OF BUILDINGS – BASIC REQUIREMENTS

Sabine (1911) was the first one to study the acoustics of hall scientifically, and proposed a formula which is the backbone of present day acoustics. Following are some of the requirements for good acoustics in a hall:

- (i) Sound must reach every part of the hall with an audible level of loudness and without echoes
- (ii) The quality of the music and speech should reach everyone
- (iii) The reverberation must be optimum
- (iv) The sound should be distributed uniformly, throughout the audience. There should not be undesired focussing of sounds or dead spots in the hall
- (v) There should not be any overlapping of syllable or musical notes
- (vi) Resonance in the hall should be avoided
- (vii) Extraneous noise should not reach the hall
- (viii) Sufficient windows and ventilators
- (ix) The source should generate sound of adequate intensity. This can be achieved using loud speakers, sound reflectors, etc.

18.2 SOUND INTENSITY

The intensity of a sound wave I means the rate of flow of sound energy through unit area parallel to a wave front. As energy per unit time is power, the intensity is measured in units of W/m^2 . Though the basic unit is W/m^2 , it is the practice to specify it in a logarithmic scale. The unit is **bel** or more commonly, the **decibel** (dB) which is 0.1 bel. Often the intensity level β is measured with respect to a standard zero level intensity I_0 , which is the lower threshold of hearing equal to 10^{-12} W/m^2 at 1000 Hz frequency:

$$\beta(\text{dB}) = 10 \log \frac{I}{I_0} \quad (18.1)$$

where the logarithm is to the base 10. Thus the intensity level of a sound whose intensity $I = 10^{-5} \text{ W/m}^2$ will be

$$\begin{aligned} \beta &= 10 \log \left(\frac{10^{-5} \text{ W/m}^2}{10^{-12} \text{ W/m}^2} \right) = 10 \log 10^7 \\ &= 10 \times 7 = 70 \text{ dB} \end{aligned} \quad (18.2)$$

The intensities and intensity level for number of common sounds are listed in Table 18.1.

Table 18.1 Intensity of various types of sound

Source of the sound	Intensity level (dB)	Intensity (W/m^2)
Threshold of hearing	0	1×10^{-12}
Whisper	20	1×10^{-10}
Quiet radio	40	1×10^{-8}
Ordinary conversation, at 50 cm	65	3×10^{-6}
Busy street traffic	70	1×10^{-5}
Siren at 30 m	100	1×10^{-2}
Loud indoor rock concert	120	1
Jet plane at 30m	140	100

18.3 ABSORPTION COEFFICIENT

Different surfaces absorb sound energy differently. An open window passes all the sound waves falling on it, and hence a open window is taken as a perfect absorber. The extent of absorption by a surface is expressed in terms of the quantity absorption coefficient. As the open window is a perfect absorber, absorption coefficient of all substances are measured in terms of **open window unit** (OWU). The **absorption coefficient** α of a surface is defined as the ratio of sound energy absorbed by the surface to the sound energy absorbed by an equal area of a perfect absorber given as

$$\alpha = \frac{\text{Sound energy absorbed by the surface}}{\text{Sound energy absorbed by an equal area of a perfect absorber}}$$

Sound absorption coefficient of some materials are listed in Table 18.2.

Table 18.2 Sound absorption coefficient of certain materials at 500 Hz frequency range

Material	Absorption coefficient (OWU)
Open window	1.0
Marble	0.01
Concrete	0.17
Asbestos	0.26
Carpet	0.30
Fibre board	0.50
Heavy curtains	0.50
Fibre glass	0.75
Human body	4.5

18.4 REVERBERATION

The sound produced by a source in a hall suffers successive reflections from the wall, floor, ceiling and other reflecting materials in the hall. Hence, in addition to the sound of the theater or a song or a movie, the listeners hear a series of sound waves. This gives the listener a persistence of sound even after the original sound has ceased. This is called **reverberation**. We can feel the reverberation of sound in a hall even after the source producing the sound is turned off.

The intensity of sound produced in a hall decays exponentially to zero with time. The decay time will be less if sound absorbing materials like audience, windows, curtains, etc. are present in the hall. Consequently, in such situations the reverberation will be less.

18.4.1 Reverberation Time

The interval of time required for the intensity to drop to one millionth of its original value is called **reverberation time**. It can be expressed in terms of sound level in dB. If the incident intensity is I_i , then the final intensity I_f is one millionth of I_i . That is,

$$I_f = 10^{-6} I_i \quad \text{or} \quad \frac{I_i}{I_f} = 10^6 \quad (18.3)$$

In terms of decibels, we get

$$\begin{aligned} \text{dB}_i &= 10 \log \frac{I_i}{I \text{ (Standard)}} \\ \text{dB}_f &= 10 \log \frac{I_f}{I \text{ (Standard)}} \\ \text{dB}_i - \text{dB}_f &= 10 \log \frac{I_i}{I_f} = 10 \log 10^6 \\ &= 10 \times 6 = 60 \end{aligned} \quad (18.4)$$

In other words, reverberation time for a hall is the time required for the intensity to drop by 60 dB.

18.5 SABINE'S FORMULA FOR REVERBERATION TIME

The relation connecting reverberation time, the volume of the hall (V), the area (S) and the absorption coefficient (α) is known as the Sabine's formula. Let $\alpha_1, \alpha_2, \alpha_3, \dots$, be the absorption coefficients of the materials in the hall whose surface areas exposed to sound be s_1, s_2, s_3, \dots , respectively. Then the average value of the absorption coefficient is given as

$$\alpha = \frac{\alpha_1 s_1 + \alpha_2 s_2 + \alpha_3 s_3 + \dots}{s_1 + s_2 + s_3 + \dots} = \frac{\sum_i \alpha_i s_i}{S}$$

$$\sum_i \alpha_i s_i = \alpha S \quad (18.5)$$

where S is the total surface area of all surfaces exposed to sound. By statistical method, Jaeger has shown that sound travels an average distance of $4V/S$, where V is the volume of the hall, between two successive reflections. This distance is known as **mean free path**.

$$\text{Time taken between two successive reflections} = \frac{4V}{Sv}$$

where v is the velocity of sound.

The average number of reflections per second is

$$n = \frac{Sv}{4V} \quad (18.6)$$

Let I_t be the average intensity per unit volume of the hall at any given instant and dI the drop in intensity in a small interval of time dt . Then

The number of reflections in time $dt = ndt$

Drop in intensity per reflection = αI

Drop in intensity for ndt reflections is $dI = -\alpha I n dt$

The negative sign is to indicate the drop in intensity. Substituting the value of n , we get

$$dI = \frac{-\alpha I S v dt}{4V}$$

$$\frac{dI}{I} = -\frac{\alpha S v dt}{4V}$$

Integrating, we get

$$\ln I_t = -\frac{\alpha S v t}{4V} + \text{Constant}$$

Denoting the initial intensity by I_0 , we obtain

$$\ln\left(\frac{I_t}{I_0}\right) = -\frac{\alpha Sv t}{4V}$$

$$\frac{I_t}{I_0} = e^{-\frac{\alpha Sv t}{4V}}$$

Using Eq. (18.3), we obtain

$$e^{-\frac{\alpha Sv T}{4V}} = 10^{-6} \text{ where } T \text{ is the reverberation time.}$$

$$\Rightarrow \frac{\alpha Sv T}{4V} = 6 \ln 10 = 6 \times 2.3026$$

Taking the velocity of sound at room temperature as 330 ms^{-1} , we get

$$T = \frac{6 \times 2.3026 \times 4V}{(330 \text{ ms}^{-1}) \alpha S}$$

$$T = \frac{0.167V}{\alpha S} = \frac{0.167V}{\sum_i \alpha_i s_i} \quad (18.7)$$

This is **Sabine's formula** for reverberation time which is in fairly good agreement with experimental values. From Eq. (18.7), it is obvious that the reverberation time is:

- (i) Directly proportional to the volume of the hall
- (ii) Inversely proportional to the total absorption
- (iii) Inversely proportional to the total area of sound absorbing surfaces.

18.6 ARCHITECTURAL ACOUSTICS – FACTORS AFFECTING AND REMEDIES

In a hall that is acoustically well-designed, the sound produced at one place, whether speech or music, should reach every point in the hall with the required audibility, and then it should die away at the appropriate time. Therefore, while designing the acoustics of a hall the following factors should be taken into consideration:

18.6.1 Reverberation

For good acoustics, the reverberation time should neither be too small nor too large. If it is too small, the loudness will be inadequate and the sound dies away in a very short time. This gives the hall a dead effect. When the reverberation time is too large, the sound persists for a longer time resulting in the overlap of successive sounds. This results in loss of clarity which makes the sound unintelligible. Therefore, the reverberation time must have an appropriate value so that everyone in the hall hears the sound clearly. This optimum value is referred to as **optimum reverberation time**. The suitable reverberation time can be evaluated using Sabine's formula, given in Eq. (18.7)

$$T = \frac{0.167 V}{\alpha S}$$

Experimentally, it is found that for distinct hearing in a hall of volume 300 m^3 the optimum value of reverberation time is about 1.03 s for speech. As the frequency involved is higher for music, the time would be slightly larger. The reverberation time of a hall can be controlled by the following factors:

- (i) The first and foremost factor is having sufficient windows and ventilators. These are important as we can keep the windows open or close so that the reverberation time is optimum.
- (ii) Having audience as per the capacity of the hall, as human beings are good absorbers.
- (iii) Lining the walls and roof with good absorbers like fibre boards, glass wool, felt, etc. is another factor.
- (iv) Using thick curtains
- (v) Covering the floors with carpets
- (vi) Providing suitable sound absorbers on the walls

The presence of the above items reduces reverberation in a hall.

18.6.2 Focussing Surfaces

The presence of focussing surfaces such as concave, spherical, cylindrical, and parabolic helps in the concentration of sound in certain regions, which causes less sound or no sound in certain other regions. If reflecting surfaces are present, the reflected sound waves may combine with the direct sound waves and produce stationary waves which make the sound intensity non-uniform. Hence, for uniform distribution of sound energy:

- (i) Curved surfaces should be avoided or at least minimized
- (ii) Ceiling should be low
- (iii) A parabolic surface may be arranged with the speaker at its focus. This sends out uniform sound energy in the entire hall, as shown in Fig. 18.1.

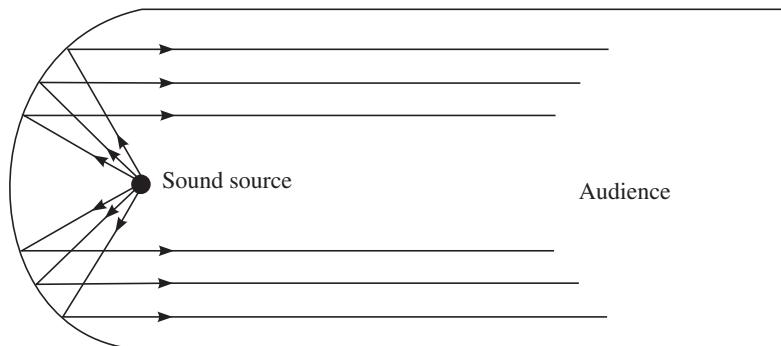


Fig. 18.1 Parabolic surface with the speaker at the focus.

18.6.3 Sufficient Loudness

An average person would be able to detect a 30 dB sound at 1000 Hz as reasonably loud. However, a 30 dB sound at 50 Hz would not be heard at all. Therefore, for satisfactory hearing, sufficient loudness throughout the hall is a necessity. The following arrangements would increase the loudness:

- (i) By setting up loud speakers at different positions in the hall, the additional sound energy required can be produced.
- (ii) Low ceiling of suitable shape can also serve the purpose by reflecting sound energy towards the audience (Fig. 18.2).
- (iii) By keeping large sounding boards behind the speaker and facing the audience.

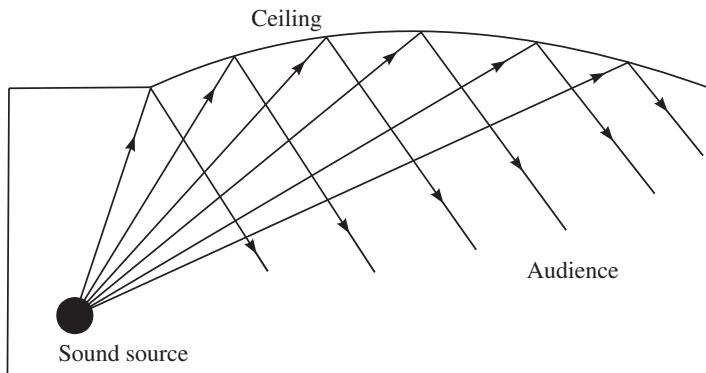


Fig. 18.2 Hall with low ceiling of suitable shape.

18.6.4 Absence of Echoes

When direct and reflected sound waves coming from the same source reach the listener in a time interval greater than $(1/7)$ s, echo is heard, which causes confusion. Hence, for good hearing, echo must be avoided. This can be achieved by covering walls at distance and the high ceiling with sound absorbers.

18.6.5 Resonance

Items including wall that are not rigid can cause forced oscillation and resonance. Due to the interference from the sound thus created, the original sound may be distorted. Locked up air can also create resonance. The resonant vibrations must be suitably damped.

18.6.6 Echelon Effect

Regular spacing of reflecting surfaces and railings may produce additional notes due to the regular succession of echoes. This makes the original sound unintelligible. Such surfaces should be avoided or covered with absorbers.

SOLVED EXAMPLES

Example 18.1 A point sound source is generating sound at a point. At a distance of 200 m away, the sound reduces to a level to 60 dB. Calculate the output power of the sound source in W/m².

Solution: Let I be the sound intensity at a distance of 200 m. For the intensity, we have

$$\text{dB} = 10 \log_{10} \frac{I}{I_0}$$

$$60 = 10 \log \frac{I}{I_0} \quad \text{or} \quad \log \frac{I}{I_0} = 6$$

$$\frac{I}{I_0} = 10^6$$

$$I = (10^{-12} \text{ W/m}^2) 10^6 = 10^{-6} \text{ W/m}^2$$

$$\text{Power of the source} = 4\pi r^2 I$$

$$= 4\pi(200 \text{ m})^2 \times 10^{-6} \text{ W/m}^2$$

$$= 0.5 \text{ W}$$

Example 18.2 If the intensity of sound in W/m² is doubled, what is the change in sound level in dB?

Solution: Let dB₁ be the sound level in dB when the sound intensity is I_1 W/m² and dB₂ is the sound level when the intensity is I_2 W/m².

$$\text{dB}_1 = 10 \log \frac{I_1}{I_0} \quad \text{and} \quad \text{dB}_2 = 10 \log \frac{I_2}{I_0}$$

$$\text{dB}_2 - \text{dB}_1 = 10 \log \frac{I_2}{I_0} - 10 \log \frac{I_1}{I_0}$$

$$= 10 \log \frac{I_2}{I_1} = 10 \log 2$$

$$= 10 \times 0.301 = 3.01$$

The sound intensity level increased by 3 dB.

Example 18.3 The reverberation time of a hall of volume 8000 m³ is 1.5 s. Calculate the total absorption in the hall.

Solution: From Sabine's formula, we have

$$T = \frac{0.167V}{\sum_i \alpha_i s_i} \quad \text{or} \quad \sum_i \alpha_i s_i = \frac{0.167V}{T}$$

$$\sum_i \alpha_i s_i = \frac{0.167 \times 8000 \text{ m}^3}{1.5 \text{ s}} = 890.7$$

The total absorption = 890.7 OWU.

Example 18.4 A hall has dimensions of $25 \text{ m} \times 20 \text{ m} \times 8 \text{ m}$. The reverberation time is 4s. Determine the average absorption coefficient of the surfaces.

Solution: Volume of the hall = $25 \times 20 \times 8 \text{ m}^3 = 4000 \text{ m}^3$

$$\begin{aligned}\text{The total area } S &= 2[(25 \times 20) + (25 \times 8) + (20 \times 8)] \text{ m}^2 \\ &= 1720 \text{ m}^2\end{aligned}$$

From Sabine's formula, we obtain

$$T = \frac{0.167V}{\alpha S} \quad \text{or} \quad \alpha = \frac{0.167V}{TS}$$

Average absorption coefficient is given by

$$\begin{aligned}\alpha &= \frac{0.167 \times 4000 \text{ m}^3}{4 \text{ s} \times 1720 \text{ m}^2} = 0.097 \text{ ms}^{-1} \\ &= 0.097 \text{ OWU m}^{-2}\end{aligned}$$

Example 18.5 In a hall the area of the floor and ceiling is 100 m^2 each. The area of wall is 200 m^2 . The absorption coefficients of the wall, ceiling and floor are 0.025, 0.02 and 0.55, respectively. If the volume of the hall is 475 m^3 , calculate the reverberation time for the hall.

Solution: From Sabine's formula, we have

$$\begin{aligned}T &= \frac{0.167V}{\sum_i \alpha_i s_i} \\ \sum_i \alpha_i s_i &= (200 \times 0.025 + 100 \times 0.02 + 100 \times 0.55) \\ &= 62\end{aligned}$$

$$\begin{aligned}\text{Reverberation time } T &= \frac{0.167 \times 475}{62} \\ &= 1.28 \text{ s.}\end{aligned}$$

Example 18.6 What is the resultant sound level when a 80 dB sound is added to a 70 dB level?

Solution: For the 80 dB sound

$$80 \text{ dB} = 10 \log \frac{I_1}{I_0} \quad \text{or} \quad 8 = \log \frac{I_1}{I_0}$$

$$I_1 = 10^8 I_0$$

For the 70 dB sound

$$I_2 = 10^7 I_0$$

Adding the two, we get

$$I = I_1 + I_2 = 10^8 I_0 + 10^7 I_0 = 11 \times 10^7 I_0$$

The intensity level of the sum in dB is

$$\begin{aligned} \text{dB} &= 10 \log \frac{I}{I_0} = 10 \log(10^7 \times 11) \\ &= 10(7 \log 10 + \log 11) = 10[7 + 1.041] \\ &= 80.041 \end{aligned}$$

The resultant intensity is 80.041 dB.

REVIEW QUESTIONS

- 18.1** Explain intensity of sound. What is its unit?
- 18.2** What is reverberation? Define reverberation time.
- 18.3** Why is an optimum value of reverberation time needed for maintaining the acoustics of a hall?
- 18.4** Define absorption coefficient of a material. What is its unit?
- 18.5** Explain the factors affecting the acoustic quality of a building.
- 18.6** State Sabine's formula for reverberation time and explain the quantities involved.
- 18.7** State Sabine's formula. How can it be used for determining the absorbing power of surfaces involved?
- 18.8** Give an account of the factors affecting the acoustics of a hall. How are they remedied?
- 18.9** Derive Sabine's formula for reverberation time and explain its importance.

PROBLEMS

1. If the sound level in dB increases from 10 dB to 20 dB, what is the ratio of increase in sound intensity in W/m²?
2. A point source emits sound of output power 50 W. Determine the distance at which sound reduces to a level of 30 dB.
3. The intensity of sound at a place is 50 dB. What should be the intensity of an additional sound source to get a resultant intensity of 60 dB?

364 *Engineering Physics*

4. A hall has a volume of 7500 m^3 . What should be the total absorption in the hall, if a reverberation time of 1.5 s is to be maintained?
5. A hall has a volume of 12500 m^3 and reverberation time of 1.5 s. If 200 cushioned chairs are additionally placed in the hall, what will be the new reverberation time of the hall? The absorption of each chair is 1.0 OWU.
6. The intensity level of sound from a jet plane at a distance of 20 m is 140 dB. What is the intensity level at 200 m?

APPENDIX



Abbreviation for Units and Physical Constants

Table A1.1 Abbreviation for units

Unit	Abbreviation	Unit	Abbreviation
Ampere	A	Metre	m
Angstrom	Å	Micrometre	µm
Coulomb	C	Microsecond	µs
Degree celsius	°C	Microcoulomb	µC
Centimetre	cm	Minute	min
Electron-volt	eV	Millimeter	mm
Degree fahrenheit	°F	Millisecond	ms
Gauss	G	Nanometre	nm
Gram	g	Ohm	Ω
Henry	H	Second	s
Hour	h	Tesla	T
Hertz	Hz	Unified mass unit	u
Joule	J	Volt	V
Kelvin	K	Watt	W
Kilogram	kg	Weber	Wb
Kilometre	km	Year	y
Litre	L		

366 Appendix I—Abbreviation for Units and Physical Constants

Table A1.2 Fundamental physical constants

Quantity	Symbol	Value
Elementary charge	e	1.60218×10^{-19} C
Speed of light in vacuum	c	2.99792×10^8 m s ⁻¹
Planck's constant	h	6.62608×10^{-34} Js
	$\hbar = \frac{h}{2\pi}$	1.05457×10^{-34} Js
Permeability of vacuum	μ_0	$4\pi \times 10^{-7}$ N A ⁻² (F/m)
Permittivity of vacuum	ϵ_0	$8.85419 \text{ C}^2\text{N}^{-1} \text{ m}^{-2}$ (H/m)
Avogadro number	N_A	6.02214×10^{23} mol ⁻¹
Boltzman constant	k = R/N _A	1.38066×10^{-23} J K ⁻¹
Molar gas constant	R = kN _A	8.31451 J mol ⁻¹ K ⁻¹
Atomic mass unit	u	1.66054×10^{-27} kg
Mass of an electron	m _e	9.10939×10^{-31} kg 0.511 MeV/c ² 5.4858×10^{-4} u
Mass of a proton	m _p	1.67262×10^{-27} kg 938.2723 MeV/c ² 1.00728 u
Mass of a neutron	m _n	1.67493×10^{-27} kg 939.5656 MeV/c ² 1.00866 u

Table A1.3 Useful combination of constants

Bohr radius	$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}$	5.2918×10^{-11} m
Coulomb constant	$\frac{1}{4\pi\epsilon_0}$	8.9876×10^9 Nm ² C ⁻²
	$\frac{e^2}{4\pi\epsilon_0}$	2.3071×10^{-28} Jm
Bohr magneton	$\mu_B = \frac{e\hbar}{2m_e}$	9.274×10^{-24} J T ⁻¹
Hydrogen ground state energy	$-\frac{e^2}{8\pi\epsilon_0 a_0}$	-13.606eV = -2.1799 × 10 ⁻¹⁸ J
Rydberg constant	$R_\infty = \frac{\alpha^2 m_e c}{2h}$	1.09737×10^7 m ⁻¹
Hydrogen Rydberg	$R_H = \frac{\mu R_\infty}{m_e}$	1.09678×10^7 m ⁻¹
Magnetic flux quantum	$\Phi_0 = \frac{\hbar}{2e}$	2.0678×10^{-15} Tm ²

Table A1.4 Conversion factors

$1 \text{ C} = 1 \text{ As}$	$1 \text{ eV} = 1.6022 \times 10^{-19} \text{ J}$
$1 \text{ cal} = 4.186 \text{ J}$	$1 \text{ \AA} = 10^{-10} \text{ m} = 10^{-8} \text{ cm} = 10^{-4} \mu\text{m}$
$1 \text{ cm}^{-1} = 100 \text{ m}^{-1}$	$1 \text{ T} = 10^4 \text{ G} = 1 \text{ N/A}$
$1 \text{ u} = 931.4943 \text{ MeV/c}^2$	

Table A1.5 Prefixes for power of ten

Power	Prefix	Symbol	Power	Prefix	Symbol
10^{18}	exa	E	10^{-1}	deci	d
10^{15}	peta	P	10^{-2}	centi	c
10^{12}	tera	T	10^{-3}	milli	m
10^9	giga	G	10^{-6}	micro	μ
10^6	mega	M	10^{-9}	nano	n
10^3	kilo	k	10^{-12}	pico	p
10^2	hecto	h	10^{-15}	femto	f
10^1	deka	da	10^{-18}	atto	a

APPENDIX



Fraunhofer Diffraction at a Single Slit

The rays diffracted from the single slit AB in the direction of incident ray are focussed at C and those diffracted with an angle θ is focussed at P . The secondary wavelets from AB reaches C in phase, because the point C is optically equidistant from all points. Hence C has maximum intensity.

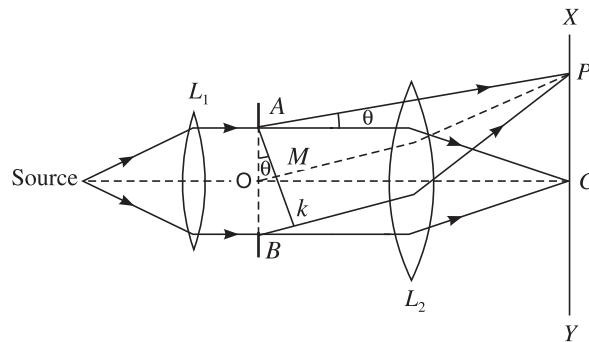


Fig. AII.1 Fraunhofer diffraction at a single slit.

In order to find the resultant intensity at P draw AK perpendicular to BK . Now, the optical length of all the rays from the plane AK to the point P are same. The path difference between the extreme points A and B is given by

$$\begin{aligned} P &= BK = AB \sin \theta \\ &= a \sin \theta \end{aligned}$$

where a is the width of AB . Hence the phase difference between the extreme points A and B

$$\begin{aligned} &= \frac{2\pi}{\lambda} \times \text{path difference} \\ &= \frac{2\pi}{\lambda} a \sin \theta \end{aligned}$$

The aperture AB can be divided into n equal number of parts, each part being a source of secondary wavelets. Even though the amplitude from different parts are same, their phase will gradually vary from zero to $\frac{2\pi}{\lambda} a \sin \theta$. Therefore the phase difference between the waves from consecutive parts is

$$\frac{1 \times 2\pi a \sin \theta}{n\lambda} = \delta$$

The resultant of n vibrations having common phase difference δ between successive vibrations with amplitude a is

$$\begin{aligned} R &= \frac{a \sin \frac{n\delta}{2}}{\sin \frac{\delta}{2}} \\ &= \frac{a \sin \left[\frac{\pi a}{\lambda} \sin \theta \right]}{\sin \left[\frac{\pi a}{n\lambda} \sin \theta \right]} = \frac{a \sin \alpha}{\sin (\alpha/n)} \end{aligned}$$

where

$$\alpha = \frac{\pi a \sin \theta}{\lambda}$$

when n is very large α/n is very small. Therefore $\sin (\alpha/n) \approx \alpha/n$. Hence

$$R = \frac{a \sin \alpha}{\alpha/n} = \frac{na \sin \alpha}{\alpha}$$

Since na = amplitude, when all the vibrations are in phase and writing

$$na = A$$

$$R = \frac{A \sin \alpha}{\alpha}$$

Thus, the resultant intensity at P is,

$$I = R^2 = \frac{A^2 \sin^2 \alpha}{\alpha^2}$$

$$I = \frac{I_0 \sin^2 \alpha}{\alpha^2}$$

APPENDIX



Metallic Glasses

Ordinary glasses are usually alloys of inorganic oxides such as SiO_2 , Na_2O , CaO , and other oxides. By cooling or quenching certain molten metallic alloys rapidly, one can produce amorphous or glassy materials called **metallic glasses**. The study of metallic glasses is important since they have very unusual and unique properties.

In ordinary glass manufacturing, the molten inorganic oxides when cooled, solidifies in two ways. If the rate of cooling is below a critical rate, a crystalline solid is formed whereas, if the rate of cooling is faster than the critical rate, it becomes a super cooled liquid which becomes a glass at a lower temperature. For metallic alloys, the formation of the glass structure is rather difficult and requires faster cooling rates.

Metallic glasses are prepared by combining either some metals with other metals or metalloids. Usually metals like Fe, Ni, Co, Al etc., are combined with metalloids B, C, Si, P, Ge, and As with cooling rates of the order of 10⁶ K/s. Although a number of cooling methods are available, the one which is most commonly used is known as **melt spinning** or **splatt quenching technique**. By injecting a narrow stream of liquid alloy on to a rotating cold wheel with the above cooling rate, metglass ribbons such as $\text{Fe}_{80}\text{B}_{20}$, $\text{Ni}_{40}\text{Fe}_{40}\text{P}_{14}\text{B}_6$, and $\text{Cu}_{50}\text{Zr}_{50}$ are produced. Amorphous thin films can be produced by sputtering deposition. It is also discovered that when a mixture of two elements is annealed at a temperature just high enough to promote sufficient atomic diffusion, an amorphous phase can result due to inter-diffusion.

The atomic structure of metallic glasses can be studied by diffraction experiments using x-rays, neutron diffraction and electron diffraction. A typical diffraction pattern consists of several broad peaks rather than a number of sharp Bragg peaks as for crystalline solids.

AIII.1 PROPERTIES OF METALLIC GLASSES

Metallic glasses have the following characteristics:

- (i) Better corrosion resistance than their crystalline counterpart.
- (ii) Very high mechanical strength, often of the order of 300 kg/mm² or more, which is comparable to or higher than that of the strongest steel (piano) wires.
- (iii) Exhibit considerable ductility in compression.
- (iv) Show relatively high electrical resistivity ranging from 50 to 200 $\mu\Omega$ cm. Resistivity is weakly dependent on temperature, with the temperature coefficient being either positive or negative.
- (v) Some of the glassy metals containing Zr, Mo, and La are superconducting at low temperatures.
- (vi) Very high strength coupled with ductility.
- (vii) Smaller hysteresis loss, hence they are soft ferromagnets.
- (viii) Better chemical homogeneity.
- (ix) Less dislocations.
- (x) Absence of grain boundaries.

AIII.2 APPLICATIONS

Because of the varied properties, glassy metals are used for different purposes. Some of the important applications are listed below:

- (i) Used in power transformers resulting in significant energy savings.
- (ii) In switching regulators which supply stable d.c power to computers and other devices, they save much space and weight.
- (iii) Sputter deposited Fe-Co-Tb thin films are extensively used for magneto-optical data storage.
- (iv) Used in anti-theft devices.

APPENDIX

IV

Shape Memory Alloys

Shape memory alloys (SMA) are metals which when deformed at some relatively low temperature, upon exposure to some higher temperature will return to their original shape. In most SMA's a temperature change of only about 10°C is necessary to initiate the change. Materials that exhibit shape memory only upon heating are said to have a one way shape memory, while those which undergo a change in shape on recooling have a two way memory. Typical and widely used alloys include NiTi, CuZnAl, CuAlNi, FePt and AuCd. This unusual property is due to a solid state phase change that occurs in the material.

The two phases that occur in an SMA are martensite and austenite. Martensite is the relatively soft and easily deformable phase that exists at lower temperatures. On heating, this phase attains the second form, the austenite, the stronger phase of SMA's, which occurs at higher temperatures. The shape memory effect is observed when the temperature of SMA is cooled to below the transformation temperature. At this stage the alloy is completely composed of twinned martensite which can easily be deformed. After deforming the SMA, the original shape can be recovered simply by heating the wire. The heat supplied effects the molecular rearrangement of the alloy. The deformed martensite is now transformed to the cubic austenite phase. Figure AIV.1 shows the sequences of the shape memory effect.

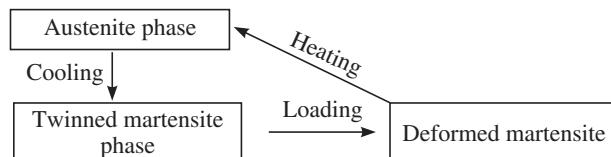


Fig. AIV.1 Schematic diagram of shape memory effect.

The systems that have reached commercial exploitation are nickel–titanium alloys and the copper base alloys. The properties of the two systems are quite different. The nickel–titanium alloys have greater shape memory strain, tend to be much more thermally stable, have excellent corrosion resistance, and have higher ductility. The copper base alloys are cheaper and have a wider range of potential transformation temperatures. The critical temperature for NiTi is approximately 250 °C and for CuZnAl approximately 90 °C. One has to be cautious while joining SMAs to conventional materials since they undergo expansions and contractions not encountered in traditional materials.

AIV.1 APPLICATIONS

Some of the applications of shape memory alloys are indicated as follows:

Blood Clot Filter

In a blood clot filter a NiTi wire is shaped to anchor itself in a vein and catch passing clots. The part is chilled so it can be collapsed and inserted into the vein, then body heat is sufficient to return the part to its functional shape.

Fluid Flow Device

Another application is a device for controlling the rate of fluid flow by carefully heating a shape memory alloy component just enough to close a valve by the desired amount.

Electrical Connector Systems

In the connector system, shape memory alloy component is used to force open a spring when the connector is heated. This allows force-free insertion or withdrawal of a circuit board in the connector.

AIV.2 OTHER APPLICATIONS

- (i) Eye glass frames use superelastic NiTi alloy to absorb large deformations without damage.
- (ii) Guide wires for steering catheters into vessels in the body have been developed using NiTi alloy wire which resists permanent deformations.
- (iii) Shape memory alloys have found application in the field of robotics.
- (iv) NiTi SMAs have also been employed in artificial joints. These alloys have also been used for bone plates, for marrow pins for healing bone fractures, and for connecting broken bones.
- (v) Used in electrical switches, actuators, temperature fuses, and fire alarms.

APPENDIX

V

Biomaterials

A biomaterial is defined as any pharmacologically inert substance or combination of substances utilized for implantation within or incorporation with a living system to supplement or replace functions of living tissues or organs. A biomedical implant is one that has the following properties:

- (i) Sufficient bonding with the surrounding body tissues
- (ii) The required mechanical characteristics such as modulus of elasticity, toughness, ductility, bending strength, and flexural strength
- (iii) The required electrical characteristics such as resistivity, piezoelectricity, etc.

Among metals, titanium – aluminum, stainless steel, and cobalt – chromium alloys are the three main alloys universally accepted for most of the high load bearing applications in skeletal system. For electrical stimulation of the heart, muscles, and nervous tissues, conducting metals like platinum and platinum iridium alloys are used. Nitinol, an alloy of nickel and titanium, finds application in orthodontics. High purity alumina is highly corrosion and wear resistant. It is also an inert material. Due to these reasons, it is often preferred over metals as an orthopedic joint replacement.

The carbons are inert ceramic materials which show varied and unique properties. In the quasi-crystalline forms, the arrangements of the crystallites and pores are important in determining the properties of carbons. The carbons used in medical devices are the quasi-crystalline ones. Carbon has good biocompatibility with bone tissues and other tissues. It also has a high strength and an elastic modulus close to that of the bone. Carbon coatings find wide applications in heart valves, blood vessel grafts, percutaneous devices because of exceptional compatibility with soft tissues and blood. The degradable ceramics which are mainly based on calcium phosphates find applications in hard tissue regeneration.

The physical properties of polymers closely resemble those of soft tissues. Therefore, they are used extensively to replace the functions of soft tissues like skin, tendons, cartilage, vessel walls, breast, and bladder. Some of the polymers used as biomaterials include polyolefins, polyamides, polyesters, polyurethanes, polyacrylates, polyethers, and silicone rubbers. Some of these are used as sutures, tissue adhesives, catheters, and space fillers.

Reinforcement by particles and fibres improves the properties of biomaterials. To improve the strength, the rubber used in catheters is reinforced with fine particles of silica (SiO_2). Composite materials used for dental purposes are a blend of glass or ceramic particles dispersed in a polymeric organic matrix. In recent years nanomaterials are also being used as biomedical implants.



Bibliography

- Aruldas, G. and Rajagopal P. *Modern Physics*, Prentice-Hall of India, New Delhi, 2006.
- Aruldas, G., *Classical Mechanics*, Prentice-Hall of India, New Delhi, 2008.
- Aruldas, G., *Quantum Mechanics*, 2nd ed., PHI Learning, New Delhi, 2009.
- Christman, J. Richard, *Fundamentals of Solid State Physics*, John Wiley and Sons Inc., New York, 1988.
- Edward, C. Jordan and Keith, G. Balmain, *Electromagnetic Waves and Radiating Systems*, Prentice-Hall of India, New Delhi , 1991.
- Jurgen, R. Meyer-Arendt., *Introduction to Classical and Modern Optics*, 2nd ed., Prentice-Hall of India, New Delhi, 1988.
- Kittel Charles, *Introduction to Solid State Physics*, Wiley Eastern, New Delhi, 1974.

Answers to Problems

Chapter 1

1. 0.158 N

3. $m = 0.52 \text{ kg}$

5. $t = 0.33 \text{ s}$

7. $1 \text{ cm}; \quad 4 \text{ cm}; \quad 100 \text{ s}^{-1}; \quad 0.01 \text{ s}; \quad 400 \text{ cm s}^{-1}$

9. $\sqrt{3}$

2. (i) 5.74° (ii) 11.54° (iii) $10 \pi \text{ s}$

4. $545.5 \text{ m}; \quad 187.5 \text{ m}$

6. $2 \times 10^{-3} \text{ m}; \quad 5 \text{ m}; \quad 50 \text{ s}^{-1}; \quad 0.02 \text{ s}; \quad 250 \text{ ms}^{-1}$

8. $y = 0.02 \sin 10\pi \left(\frac{x}{3} - 110t \right)$

10. 0.44

Chapter 2

1. $k = 0.68 \times 10^5 \text{ m}^{-1}; \quad v = 2.99 \times 10^8 \text{ m/s}; \quad \mathbf{H} = 1.33 \cos(kz - 2 \times 10^{14} t) \text{ A/m}$

3. $\mathbf{H} = \hat{\mathbf{y}} \frac{E_0 k}{\mu_0 \omega} \sin kx \cos \omega t$

6. $k = \frac{H_0}{E_0} \omega \mu$

Chapter 3

- | | |
|--|---|
| 1. 5440 Å | 2. 6783 Å |
| 3. 5000 Å | 4. 0.1 cm |
| 5. 5448 Å | 6. 2.95×10^{-4} radians |
| 7. 0.279 cm | 8. 0.255 cm |
| 9. 0.056 cm | 10. 7111 Å |
| 11. 384 nm | 12. 100 cm |
| 13. 0.393 μm | 14. 1.52 |
| 15. 1.3 | 16. 1.5 |
| 17. 589 nm | 18. 1030.7 Å |
| 19. 0.7797 cm | 20. 479 nm |
| 21. 0.025° | 22. 2.95 mm |
| 23. 1.498 | 24. 6.75 μm |
| 25. The reflected light is weak in the red region, strong in the blue-violet region, blue colour. | |

Chapter 4

- | | |
|-------------------------|--|
| 1. 499 nm | 2. $\lambda_1 = 518$ nm, $\lambda_2 = 593$ nm, $\lambda_3 = 737$ nm |
| 3. 2817 lines/cm | 4. 0.313×10^{-5} m; $\theta = 56.3^\circ$ |
| 5. 3 | 6. 12500 lines/cm |
| 7. 27500 | 8. 0.305 mm |
| 9. 2 mm | 10. 0.0134 mm |
| 11. $1' 18''$ | |

Chapter 5

- | | |
|------------------------------------|---------------------------------------|
| 1. 30.5° | 2. 1.23 |
| 3. 25% | 4. 32.38×10^{-5} m |
| 5. 30° | 6. 61.2° , 28.8° |
| 7. 8.61×10^{-5} cm | 8. 3.64×10^{-3} cm |
| 9. 0.824π | 10. 2λ |

Chapter 6

- | | |
|---|-------------------------------------|
| 1. (326) | 2. (340) |
| 3. 1a, 1b, 2c | 4. (i) 0.31 nm (ii) 0.117 nm |
| 5. (102); 0.19 nm | 6. 0.315 nm |
| 7. $\lambda = 0.7853$ Å; 18.23° | 8. 14.03° |
| 9. 0.352 nm | 10. 0.79 Å, 1.26 Å |

- 11.** 87.444° **12.** 17.2°
13. 68% **14.** $0.155r$
15. 1.538 \AA **16.** 39.8 pm
17. 10.8°

Chapter 7

- 1.** 0.0385 T **2.** $4.04 \times 10^4 \text{ A/m}$
3. 1.09 K **4.** (i) $10.29 \times 10^{-4} \text{ eV}$ (ii) $1.21 \times 10^{-3} \text{ m}$
5. 391.2 \AA **6.** $7.35 \times 10^{-3} \text{ m}$

Chapter 8

- 1.** 0.846 c **2.** 0.8 c
3. 0.2 c **4.** 120.8 m
5. 0.87 c **6.** $1626.4 \frac{\text{MeV}}{c}$
7. $1.67 \times 10^{-6} \text{ s}$ **8.** 0.996 c
9. 939.4 MeV **10.** 0.5 c
11. $1.44 \times 10^{-11} \text{ J}$ **12.** 0.866 c
13. 54.62 m **14.** 0.866 c
15. 0.866 c

Chapter 9

- 1.** 937.5 V **2.** (i) 0.0183 nm (ii) $4.27 \times 10^{-4} \text{ nm}$
3. 150 V **4.** 1836
5. 2.18 % **6.** $3.18 \times 10^{-14} \text{ m}$
7. $2.88 \times 10^3 \text{ ms}^{-1}$ **8.** $P(x) = \left(\frac{m\omega}{\pi\hbar} \right)^{\frac{1}{2}} e^{-\frac{m\omega x^2}{\hbar}}$
9. $\langle x \rangle = \frac{a}{2}; \quad \langle p_x \rangle = 0$ **10.** 112.1 eV
11. $L_x = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right); \quad L_y = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right); \quad L_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$
12. No

Chapter 10

1. 2.572

3. $\frac{30\pi^2\hbar^2}{mL^2}$

5. (i) 5.54 eV (ii) 3.32 eV

7. $2.54 \times 10^{28} \text{ m}^{-3}$, $E_F = 3.16 \text{ eV}$

2. 3.83

4. $\frac{3\pi^2\hbar^2}{ml^2}$

6. 7.03 eV

8. $0.152 \text{ JK}^{-1} \text{ mol}^{-1}$ **Chapter 11**1. 79.59×10^{14} 3. $3.88 \times 10^{-10} \text{ m}^2$, $6.44 \times 10^6 \text{ W/m}^2$ 5. $1.58 \times 10^{13} \text{ W/m}^2$ 2. $1.06 \times 10^{-19} \text{ m}$ 4. 7×10^{-12} **Chapter 12**1. (i) 0.291 (ii) 33.88° 3. 0.28; 32.52°

5. 1.47

7. 5.55

9. 379.5 μW 2. (i) 0.292 (ii) 33.96° (iii) 0.0214. (i) 0.248 (ii) 28.8°

6. 0.344

8. 0.796 μW **Chapter 13**

1. 62.7 V

3. $\alpha = 4.52 \times 10^{-42} \text{ Fm}^2$ 2. $\alpha_e = 2.243 \times 10^{-41} \text{ Fm}^2$ 4. $\varepsilon_r = 1.000564$ **Chapter 14**1. $8.69 \times 10^{-27} \text{ Am}^2$ 3. $32 \times 10^{-4} \text{ Am}^2$

5. 0.0126 T

2. 0.1257 T, 78 Am^{-1}

4. 0.201

6. 7.639×10^4 ; 76.39 Am^{-1} **Chapter 15**1. $72.895 \times 10^{11} \text{ s}^{-1}$ 2. $1.67 \text{ J mol}^{-1} \text{ K}^{-1}$; $21.87 \times 10^{11} \text{ s}^{-1}$ 3. $C_v(\text{B}) = 0.0864 \text{ J mol}^{-1} \text{ K}^{-1}$; $C_v(\text{A}) = 0.0032 \text{ J mol}^{-1} \text{ K}^{-1}$

4. 3.58 K

5. $0.254 \text{ J mol}^{-1} \text{ K}^{-1}$

6. $\frac{\pi^2 kT}{6E_F}$

Chapter 17

1. 2 mm

2. 68.75 kHz

Chapter 18

1. 10

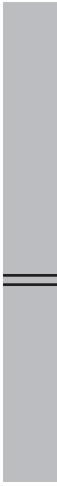
2. $6.31 \times 10^4 \text{ m}$

3. 59.542 dB

4. 835 OWU

5. 1.31 s

6. 120 dB



Index

Absence of electron in the nucleus, 190
Absorption and emission of radiation, 227
 absorption, 227
 radiative and non-radiative transitions, 229
 spontaneous emission, 228
 stimulated emission, 228
Absorption coefficient, 355
Acceptance angle and acceptance cone, 261
Acoustic grating, 348
Acoustic of buildings—basic requirements, 354
Analysis of light of unknown polarization, 107
Antiferromagnetism, 303
Antireflection coatings, 70
Atom-early models, 184
Atomic packing factor, 124

Basic postulates of quantum statistics, 213
BCS theory, 147
 cooper pairs, 147
 superconducting energy gap, 148
Bose-Einstein statistics, 213–215
Bravais lattices, 118
Breakdown of classical physics, 180
 blackbody radiation curves, 181
 optical spectra, 181
 photoelectric effect, 181
 quantization of matter and charge, 181
 specific heat, 181

Brewster's law, 101

Carbon clusters, 332
 synthesis, 332
Carbon dioxide laser, 241
Carbon nanotubes, 333
Carbon nanotubes-fabrication, 334
 chemical vapour deposition, 335
 laser evaporation method, 334
Chemical vapour deposition, 330
Classical statistics, 212
Classification of magnetic materials, 295
Clausius-Mossotti relation, 281
Coherence, 55
Comparison of the three statistics, 215
Coordination number, 122
Crystal planes and directions, 127
Curl of a vector, 33
 conduction current, 35
 displacement current, 36
 important vector identities, 34
 integral theorems, 34
 physical significance, 33

Damped harmonic oscillation, 5
Davisson-Germer experiment, 187
De Broglie wavelength, 186

- Debye model, 315
- Del operator, 30
- Density of states, 216
 - number of eigenstates in an energy range, 217
 - phase volume and eigenstates, 217
- Detection of ultrasonic waves, 346
- Diamagnetism, 296
- Dielectric loss, 283
- Dielectric polarization, 276
 - electronic polarization, 276
 - ionic polarization, 277
 - orientation polarization, 278
 - space charge polarization, 278
 - total polarization, 279
- Dielectric—Review of basic ideas, 274
 - dielectric constant, 275
 - dielectric susceptibility, 275
 - electric dipole, 274
 - polarizability, 274
 - polarization vector, 275
- Diffraction at a circular aperture, 82
- Diffraction grating, 83
 - dispersive power, 86
 - experimental arrangement, 84
 - grating equation, 85
 - transmission and reflection gratings, 83
- Diffraction phenomenon, 78
- Distribution laws of the different statistics, 214
 - Bose-Einstein statistics, 215
 - Fermi-Dirac statistics, 215
 - Maxwell-Boltzmann statistics, 214
- Divergence of a vector, 31
 - physical significance, 32
- Double refraction, 102
 - as an analyzer, 105
 - uniaxial and biaxial crystals, 103
- Dual nature of radiation, 184

- Eigenvalues and eigenfunctions of an operator, 196
 - Schrödinger equation as an eigenvalue equation, 197
- Einstein model, 312
 - high temperature limit, 313
 - low temperature limit, 314
- Einstein's coefficients, 229
- Electromagnetism and Galilean transformation, 157
 - ether hypothesis, 157
- Energy of a Fermi-Dirac system, 219
- Energy transport in wave motion, 16
- Equation of continuity for time varying fields, 36

- Faraday effect, 109
- Fermi-Dirac statistics, 214–216
- Ferrimagnetism, 305
 - ferrites, 305
- Ferroelectric crystals, 285
- Ferromagnetism, 298
 - curie temperature, 299
 - exchange interaction, 298
- Fibre losses and transmission windows, 264
- Fibre optic sensors, 268
 - intensity modulated sensor, 269
 - sensors in health care, 270
- Forced oscillation and resonance, 8
- Frauhofer diffraction at a single slit, 80
 - experimental arrangement, 80
 - positions of maximum and minimum intensities, 80
- Frequency dependence of polarization, 282

- Galilean transformation, 156
- General form of a travelling wave, 12
- General wave equation, 17
 - one-dimensional, 17
 - three-dimensional, 18
- Gradient of a scalar function, 31
 - physical significance, 31
- Grating and prism spectra-comparison, 88

- Half wave plate, 106
- Hamiltonian operator, 196
- Harmonic oscillation, 1
 - energy in SHM, 4
 - equation of motion, 2
 - simple harmonic motion (SHM), 1
 - solution of equation of motion, 2
- Harmonic wave, 13
- Heat capacity of electrons, 318
- He-Ne laser, 238
- High temperature superconductivity, 150
 - generation and transmission of electricity, 151
 - squid, 151
 - superconducting magnets, 151
- Holograms—types, 256
 - diffuse illumination of the object, 256
 - in-line and off-axis holography, 256
 - plane and volume holograms, 257
 - reflection hologram, 256
- Holography
 - application, 257

- data storage, 257
- holographic interferometry, 257
- holographic optical element, 257
- holographic microscopy, 258
- an introduction, 254
- pattern recognition, 258
- Holography—basic description, 255
 - recording of transmission hologram, 255
 - reconstruction, 255
- Huygens' theory of double refraction, 104
- Hysteresis loop, 301

- Induced birefringence, 109
- Inertial and non-inertial reference frames, 155
- Interference filter, 69
- Interference in thin films, 58
 - reflected light, 58
 - transmitted light, 60
- Internal field in a dielectric, 279
- Interplanar spacing, 128
- Isotope effect, 147

- Josephson junction, 149

- Kerr effect, 108

- Laser—applications
 - lasers in industry, 245
 - lasers in scientific research, 246
 - lasers in medical field, 248
- Laser—basic components, 232
 - active medium, 233
 - pumping methods, 233
 - resonant cavity, 234
- Laser induced evaporation method, 327
- Laser light—characteristics, 244
 - coherence, 244
 - directionality, 244
 - intensity, 245
 - monochromaticity, 244
- Laser—principle, 229
- Laser pulse method, 328
- Lattice points per unit cell in cubic lattices, 122
 - density and number of atoms per unit cell, 122
- Lattice with two atoms per primitive cell, 319
- Length contraction, 164

- Light pipe, 268
- Light propagation in step index fibre, 259
- Light propagation in variable index fibre, 262
 - different modes in a fibre, 261
- Lorentz transformation, 161

- Magnetic materials for information storage, 306
- Magnetic properties, 143
- Magnetic flux quantization, 148
- Malus law, 102
- Mass and energy, 169
- Mass in relativity, 166
- Massless particles, 170
- Matter wave, 185
 - De Broglie hypothesis, 185
 - wave packet, 186
- Maxwell's equations, 37
 - electric flux equation, 37
 - equation for induced electric field, 38
 - equation for induced magnetic field, 38
 - magnetic flux equation, 38
- Maxwell-Boltzmann statistics, 213, 214, 216
- Medical applications of nanotechnology, 341
- Medical applications of ultrasonics, 350
 - echocardiography, 351
 - ultrasound guided technique, 351
 - ultrasound imaging, 350
- Meissner effect, 144
- Metal nanoclusters-RF plasma method 329
- Metastable states, 235
- Methods to produce interference pattern, 58
- Michelson interferometer, 66
 - application, 67
 - experimental arrangement, 66
 - principle, 66
 - working, 67
- Michelson-Morley experiment, 158
 - experiment, 159
 - interferometer, 158
 - results, 160
- Millers indices, 127
 - significant features of, 128
- Modes of propagation, 262

- Nanomaterials—synthesis, 327
- Nanotechnology
 - introduction, 326
 - medical applications, 341

- Nanowires and nanomaterials
 - application, 338
 - carbon nanotubes, 337
 - giant magnetoresistance, 337
 - metal nanoclusters, 336
 - nanoparticles of semiconducting materials, 336
- Nature of light, 97
- Nearest neighbour distance, 123
- Newton's rings by reflected light, 63
 - experimental arrangement, 63
- Newton's rings with liquid film, 65
 - ring system in transmitted light, 66
- Nicol prism, 104
- Non-destructive testing, 348
 - pulse echo method-flaw detection, 348
 - sonar, 350
 - thickness gauging, 350
 - transmission system, 349
- Normalization condition, 195
- Numerical aperture, 260
- Operators, 195
- Optical communication, 266
 - communication system, 266
 - advantages of optical communication, 267
- Optical fibre, 258
- Optical sources for fibre optics, 265
- Origin of magnetic moment, 293
 - orbital magnetic moment, 293
 - spin magnetic moment, 294
 - total magnetic moment, 294
- Particle in a one-dimensional infinite square well, 199
 - energy eigenvalues and eigenfunctions, 200
 - probability distribution, 201
- Permeability and susceptibility, 292
- Phonons, 320
- Photoelastic bench, 112
- Photoelasticity, 109
 - basic ideas, 109
- Piezoelectric crystals, 285
 - applications, 286
- Planck's blackbody formula, 218
- Planck's quantum theory, 182
 - average energy of an oscillator, 183
 - Planck's hypothesis, 182
 - Planck's radiation formula, 183
- Plane and spherical waves, 21
- Point and space groups, 120
- Polarization, 96
- Polarization by reflection, 101
 - Brewster's law, 101
- Polarization by selective absorption, 107
 - dichroism, 107
 - polaroid, 107
- Ponting's theorem, 44
- Population inversion, 232
- Postulates of quantum mechanics, 197
- Postulates of special theory of relativity, 161
- Prediction of electromagnetic waves, 40
- Principle of laser, 229
- Probability interpretation of the wave function, 194
- Production of different types of polarized light, 106
- Production of ultrasonics-magnetostriction method, 343
 - experimental arrangement, 344
 - principle, 343
- Production of ultrasonics-piezoelectric method, 345
 - experimental arrangement, 344
 - principle, 343
- Properties of ultrasonics, 347
- Quality factor, 10
- Quantum mechanical tunneling, 201
- Quantum statistics, 213
- Quarter wave plate, 105
- Relation between E and H , 42
- Resolving power, 86
 - Rayleigh's criterion, 86
- Resolving power of a grating, 87
- Resolving power of a microscope, 90
- Resolving power of a telescope, 89
- Retardation plates, 105
- Reverberation, 356
 - time, 356
- Ruby laser, 236
 - applications, 238
- Sabine's formula for reverberation time, 357
- Semiconductor laser, 242
- Simultaneity, 162

- Single and multimode fibres, 263
 - salient features of a single mode fibre, 263
 - salient features of multimode fibre, 263
- Sodium chloride – structure, 126
- Soft and hard magnetic materials, 302
- Sol–gel method, 331
- Solution of the one-dimensional wave equation, 18
- Solution of the three-dimensional wave equation, 19
- Sound intensity, 355
- Space lattice, basis and translation vectors, 117
- Specific heat, 311
- Statistical methods, 210
 - cells in phase space, 212
 - macroscopic and microscopic systems, 210
 - macrostate and microstate, 211
 - phase space, 211
- Stress on a photoelastic material—An analytical treatment, 110
 - isoclinic and isochromatic fringes, 111
- Superconductivity, 142
- Superconductors
 - type-I and type-II, 145
- Superposition of waves, 50
 - analytical treatment, 51
- Synthesis of nanomaterials—General ideas, 327
- Thermal conductivity, 321
- Thermolysis, 329
- Thickness of transparent, 56
- Thomson's experiment, 188
- Three and four level laser systems, 235, 236
 - Three level system, 235
 - Four level system, 236
- Time dependent schrodinger equation, 191
 - energy and momentum operators, 193
 - free particle wave equation in one-dimension, 191
 - general wave equation, 193
- Time dilation, 163
- Time independent Schrodinger equation, 193
- Transverse nature of electromagnetic waves, 41
- Transverse vibrations or a stretched siring, 21
- Types of diffraction, 79
- Types of polarized light, 98
 - circularly polarized light, 99
 - elliptically polarized light, 99
 - linearly polarized light, 98
- Types of waves, 11
 - electromagnetic waves, 12
 - mechanical waves, 12
 - other types of waves, 12
- Ultrasonics
 - detection, 346
 - medical applications, 350
 - other applications, 351
 - properties, 347
 - production, 343–346
 - ultrasonic waves, 343
- Uncertainty in frequency of light emitted by an atom, 190
- Uncertainty principle, 189
- Unit cell, 121
 - non-primitive unit cell, 122
 - primitive unit cell, 121
- V-number of a fibre, 264
- Wave function, 190
- Wave velocity and particle velocity, 15
- Wedge shaped films, 61
 - the diameter of a thin wire, 62
 - testing the planeness of surfaces, 62
- Weiss theory of ferromagnetism, 300
- White light fringes 55
- X-ray diffraction, 130, 131
 - Bragg's law, 130
 - Laue's method, 132
 - powder method, 133
 - unit cell dimensions, 132
- Young's double slit experiment, 53
 - conditions for sustained interference pattern, 55
 - experimental arrangement, 53
 - expression for bandwidth, 53

ENGINEERING PHYSICS

G. ARULDHAS

Primarily intended for the undergraduate students of all branches of engineering, this textbook provides a sound understanding of the fundamental concepts and principles of physics in a simple and easy-to-understand language.

Organized in 18 chapters, the book exposes students to the fundamentals of oscillations and waves, interference of light, diffraction, polarization, optical instruments, laser, fibre optics, mechanics and special theory of relativity. Apart from giving a detailed theoretical analysis of these topics, it also provides a deep insight on various advanced topics such as acoustics, ultrasonics and nanotechnology, along with their applications. The pedagogical aids such as solved numerical problems and review questions are also included at the end of each chapter.

KEY FEATURES

- ◆ Numerous solved examples to stress on the conceptual understanding
- ◆ Chapter-end model questions to probe a student's grasp of the subject matter
- ◆ Chapter-end objective type questions (with answers) for self-evaluation by the students

ABOUT THE AUTHOR

Late Dr. G. ARULDHAS was former Professor and Head, Department of Physics, and former Dean, Faculty of Science, University of Kerala. He was also a CSIR Emeritus Scientist, and Director, University Institute of Technology. Dr. Aruldhas is credited with about 145 research papers in various national and international journals. He authored many books such as *Molecular Structure and Spectroscopy* (2nd ed.), *Classical Mechanics, Modern Physics* and *Quantum Mechanics* (2nd ed.) (all published by PHI Learning).

Our other useful books

Engineering Chemistry, 2nd ed., N. Krishnamurthy,
P. Vallinayagam and D. Madhavan

Engineering Mathematics, Vol. 1, 4th ed., S.S. Sastry

Engineering Mathematics, Vol. 2, 4th ed., S.S. Sastry

Communication Skills for Professionals, Nira Konar

Rs. 250.00

www.phindia.com

