

Eastern
Economy
Edition

Materials Science and Engineering

A First Course

FIFTH EDITION



V. Raghavan



THE PERIODIC TABLE

1. The atomic number is shown at the top.
2. The middle line indicates the atomic mass in amu
3. The outer electron configuration is shown at the bottom

1																			2
H 1.008 1s ¹																			He 4.003 1s ²
3	4																		
Li 6.939 2s ¹	Be 9.012 2s ²																		
11	12																		
Na 22.99 3s ¹	Mg 24.31 3s ²																		
19	20	21	22	23	24	25	26	27	28	29	30								
K 39.10 4s ¹	Ca 40.08 4s ²	Sc 44.96 3d ¹ 4s ²	Ti 47.90 3d ² 4s ²	V 50.94 3d ³ 4s ²	Cr 52.00 3d ⁵ 4s ¹	Mn 54.94 3d ⁵ 4s ²	Fe 55.85 3d ⁶ 4s ²	Co 58.93 3d ⁷ 4s ²	Ni 58.71 3d ⁸ 4s ²	Cu 63.54 3d ¹⁰ 4s ¹	Zn 65.37 3d ¹⁰ 4s ²								
37	38	39	40	41	42	43	44	45	46	47	48								
Rb 85.47 5s ¹	Sr 87.62 5s ²	Y 88.91 4d ¹ 5s ²	Zr 91.22 4d ² 5s ²	Nb 92.91 4d ⁴ 5s ¹	Mo 95.94 4d ⁵ 5s ¹	Tc 99 4d ⁵ 5s ¹	Ru 101.1 4d ⁷ 5s ¹	Rh 102.9 4d ⁸ 5s ¹	Pd 106.4 4d ¹⁰	Ag 107.9 4d ¹⁰ 5s ¹	Cd 112.4 4d ¹⁰ 5s ²								
55	56	57	72	73	74	75	76	77	78	79	80								
Cs 132.9 6s ¹	Ba 137.3 6s ²	La 138.9 5d ¹ 6s ²	Hf 178.5 5d ² 6s ²	Ta 180.9 5d ³ 6s ²	W 183.9 5d ⁴ 6s ²	Re 186.2 5d ⁵ 6s ²	Os 190.2 5d ⁶ 6s ²	Ir 192.2 5d ⁷	Pt 195.1 5d ⁹ 6s ¹	Au 197.0 5d ¹⁰ 6s ¹	Hg 200.6 5d ¹⁰ 6s ²								
87	88	89																	
Fr 223 7s ¹	Ra 226 7s ²	Ac 6d ¹ 7s ²																	
			58	59	60	61	62	63	64	65									
			Ce 140.1 4f ² 6s ²	Pr 140.9 4f ³ 6s ²	Nd 144.3 4f ⁴ 6s ²	Pm 147 4f ⁵ 6s ²	Sm 150.3 4f ⁶ 6s ²	Eu 152.0 4f ⁷ 6s ²	Gd 157.3 4f ⁷ 5d ¹ 6s ²	Tb 158.9 4f ⁹ 5d ¹ 6s ²	Dy 162.5 4f ¹⁰ 6s ²								
			90	91	92	93	94	95	96	97	98	99	100	101	102				
			Th 232 6d ² 7s ²	Pa 231 5f ² 6d ¹ 7s ²	U 238 5f ³ 6d ¹ 7s ²	Np 237 5f ⁷ 7s ²	Pu 244 5f ⁶ 7s ²	Am 243 5f ⁷ 7s ²	Cm 247 5f ⁷ 6d ¹ 7s ²	Bk 247 5f ⁸ 6d ¹ 7s ²	Cf 251 5f ¹⁰ 7s ²	Es 254 5f ¹¹ 7s ²	Fm 257 5f ⁹ 7s ²	Md 256 5f ¹³ 7s ²	No 254 5f ¹⁴ 7s ²	Lw 257 5f ¹⁴ 6d ¹ 7s ²			

Materials Science and Engineering

A First Course

FIFTH EDITION

V. Raghavan

*Formerly Professor of Materials Science
Indian Institute of Technology Delhi*

PHI Learning Private Limited

New Delhi-110001

2011

MATERIALS SCIENCE AND ENGINEERING: A First Course, 5th ed.

V. Raghavan

© 2004 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.

Cover micrograph: Martensitic plates in austenite grains with the eutectic mixture at the grain boundaries. (*Courtesy:* ASM International, Ohio.)

ISBN-978-81-203-2455-8

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

Fortieth Printing (Fifth Edition)

...

...

September, 2011

Published by Asoke K. Ghosh, PHI Learning Private Limited, M-97, Connaught Circus, New Delhi-110001 and Printed by Mohan Makhijani at Rekha Printers Private Limited, New Delhi-110020.

Contents

<i>Preface</i>	<i>ix</i>
<i>Preface to the First Edition</i>	<i>x</i>
<i>SI Units</i>	<i>xi</i>
<i>Physical Constants</i>	<i>xiv</i>
<i>Conversion Factors</i>	<i>xv</i>
<i>Greek Alphabets</i>	<i>xvii</i>
1 Introduction	1–8
1.1 Materials Science and Engineering	1
1.2 Classification of Engineering Materials	2
1.3 Levels of Structure	3
1.4 Structure–Property Relationships in Materials	7
<i>Suggestions for Further Reading</i>	8
2 Equilibrium and Kinetics	9–22
2.1 Stability and Metastability	10
2.2 Basic Thermodynamic Functions	12
2.3 The Statistical Nature of Entropy	14
2.4 The Kinetics of Thermally Activated Processes	16
<i>Summary</i>	18
<i>Problems</i>	19
<i>Multiple Choice Questions</i>	20
<i>Sources for Experimental Data</i>	22
<i>Suggestions for Further Reading</i>	22
3 Crystal Geometry and Structure Determination	23–52
GEOMETRY OF CRYSTALS	24
3.1 The Space Lattices	24
3.2 Space Lattices and Crystal Structures	29
3.3 Crystal Directions and Planes	31
STRUCTURE DETERMINATION BY X-RAY DIFFRACTION	37
3.4 The Bragg Law of X-ray Diffraction	37
3.5 The Powder Method	41
3.6 Structure Determination	42

Summary 45
Problems 46
Multiple Choice Questions 49
Sources for Experimental Data 52
Suggestions for Further Reading 52

4 Atomic Structure and Chemical Bonding **53–80**

STRUCTURE OF THE ATOM 54
4.1 The Quantum States 54
4.2 The Periodic Table 55
4.3 Ionization Potential, Electron Affinity and Electronegativity 60
CHEMICAL BONDING 63
4.4 Bond Energy, Bond Type and Bond Length 63
4.5 Ionic Bonding 65
4.6 Covalent Bonding 68
4.7 Metallic Bonding 71
4.8 Secondary Bonding 72
4.9 Variation in Bonding Character and Properties 74
Summary 76
Problems 77
Multiple Choice Questions 78
Suggestions for Further Reading 80

5 Structure of Solids **81–119**

5.1 The Crystalline and the Noncrystalline States 82
INORGANIC SOLIDS 83
5.2 Covalent Solids 83
5.3 Metals and Alloys 88
5.4 Ionic Solids 97
5.5 The Structure of Silica and the Silicates 103
POLYMERS 106
5.6 Classification of Polymers 106
5.7 Structure of Long Chain Polymers 107
5.8 Crystallinity of Long Chain Polymers 111
Summary 112
Problems 113
Multiple Choice Questions 116
Sources for Experimental Data 119
Suggestions for Further Reading 119

6 Crystal Imperfections **120–147**

6.1 Point Imperfections 121
6.2 The Geometry of Dislocations 126
6.3 Other Properties of Dislocations 131
6.4 Surface Imperfections 137

<i>Summary</i>	142
<i>Problems</i>	143
<i>Multiple Choice Questions</i>	146
<i>Sources for Experimental Data</i>	147
<i>Suggestions for Further Reading</i>	147

7 Phase Diagrams **148–177**

7.1	The Phase Rule	148
7.2	Single-component Systems	150
7.3	Binary Phase Diagrams	151
7.4	Microstructural Changes during Cooling	157
7.5	The Lever Rule	159
7.6	Summary of Phase Diagram Rules	161
7.7	Some Typical Phase Diagrams	163
7.8	Other Applications of Phase Diagrams	170
	<i>Summary</i>	171
	<i>Problems</i>	171
	<i>Multiple Choice Questions</i>	175
	<i>Sources for Experimental Data</i>	177
	<i>Suggestions for Further Reading</i>	177

8 Diffusion in Solids **178–200**

8.1	Fick's Laws of Diffusion	179
8.2	Solution to Fick's Second Law	181
8.3	Applications Based on the Second Law Solution	182
8.4	The Kirkendall Effect	189
8.5	The Atomic Model of Diffusion	190
8.6	Other Diffusion Processes	193
	<i>Summary</i>	196
	<i>Problems</i>	196
	<i>Multiple Choice Questions</i>	199
	<i>Sources for Experimental Data</i>	200
	<i>Suggestions for Further Reading</i>	200

9 Phase Transformations **201–237**

9.1	Time Scale for Phase Changes	202
	NUCLEATION AND GROWTH	204
9.2	The Nucleation Kinetics	205
9.3	The Growth and the Overall Transformation Kinetics	211
	APPLICATIONS	213
9.4	Transformations in Steel	214
9.5	Precipitation Processes	220
9.6	Solidification and Crystallization	224
9.7	The Glass Transition	226
9.8	Recovery, Recrystallization and Grain Growth	228

<i>Summary</i>	232
<i>Problems</i>	232
<i>Multiple Choice Questions</i>	235
<i>Sources for Experimental Data</i>	237
<i>Suggestions for Further Reading</i>	237

10 Elastic, Anelastic and Viscoelastic Behaviour **238–259**

ELASTIC BEHAVIOUR	239
10.1 Atomic Model of Elastic Behaviour	239
10.2 The Modulus as a Parameter in Design	243
10.3 Rubber-like Elasticity	246
ANELASTIC BEHAVIOUR	249
10.4 Relaxation Processes	249
VISCOELASTIC BEHAVIOUR	252
10.5 Spring-Dashpot Models	252
<i>Summary</i>	255
<i>Problems</i>	256
<i>Multiple Choice Questions</i>	257
<i>Sources for Experimental Data</i>	258
<i>Suggestions for Further Reading</i>	259

11 Plastic Deformation and Creep in Crystalline Materials **260–297**

PLASTIC DEFORMATION	261
11.1 The Tensile Stress-Strain Curve	261
11.2 Plastic Deformation by Slip	263
11.3 The Shear Strength of Perfect and Real Crystals	266
11.4 The Stress to Move a Dislocation	269
11.5 The Effect of Temperature on the Stress to Move a Dislocation	271
11.6 Multiplication of Dislocations during Deformation	274
11.7 Work Hardening and Dynamic Recovery	275
11.8 The Effect of Grain Size on Dislocation Motion	278
11.9 The Effect of Solute Atoms on Dislocation Motion	281
11.10 The Effect of Precipitate Particles on Dislocation Motion	283
11.11 Review of Strengthening Methods	285
CREEP	287
11.12 Mechanisms of Creep	287
11.13 Creep Resistant Materials	289
<i>Summary</i>	290
<i>Problems</i>	291
<i>Multiple Choice Questions</i>	294
<i>Sources for Experimental Data</i>	297
<i>Suggestions for Further Reading</i>	297

12 Fracture **298–314**

- 12.1 Ductile Fracture 298
- 12.2 Brittle Fracture 300
- 12.3 Fracture Toughness 304
- 12.4 The Ductile-Brittle Transition 305
- 12.5 Fracture Mechanism Maps 307
- 12.6 Methods of Protection against Fracture 308
- 12.7 Fatigue Fracture 310
- Summary* 311
- Problems* 312
- Multiple Choice Questions* 313
- Suggestions for Further Reading* 314

13 Oxidation and Corrosion **315–331**

- OXIDATION 316
- 13.1 Mechanisms of Oxidation 316
- 13.2 Oxidation Resistant Materials 318
- CORROSION 319
- 13.3 The Principles of Corrosion 320
- 13.4 Protection against Corrosion 325
- Summary* 329
- Problems* 329
- Multiple Choice Questions* 330
- Sources for Experimental Data* 331
- Suggestions for Further Reading* 331

14 Conductors and Resistors **332–354**

- 14.1 The Resistivity Range 333
- 14.2 The Free Electron Theory 334
- 14.3 Conduction by Free Electrons 338
- 14.4 Conductor and Resistor Materials 341
- 14.5 Superconducting Materials 344
- Summary* 349
- Problems* 349
- Multiple Choice Questions* 351
- Sources for Experimental Data* 354
- Suggestions for Further Reading* 354

15 Semiconductors **355–392**

- 15.1 The Energy Gap in Solids 356
- 15.2 Intrinsic Semiconductors 361
- 15.3 Extrinsic Semiconductors 364
- 15.4 Semiconductor Materials 368
- 15.5 Fabrication of Integrated Circuits 369
- 15.6 Some Semiconductor Devices 379

<i>Summary</i>	384
<i>Problems</i>	385
<i>Multiple Choice Questions</i>	387
<i>Sources for Experimental Data</i>	392
<i>Suggestions for Further Reading</i>	392

16 Magnetic Materials **393–411**

16.1	Terminology and Classification	394
16.2	Magnetic Moments due to Electron Spin	395
16.3	Ferromagnetism and Related Phenomena	397
16.4	The Domain Structure	399
16.5	The Hysteresis Loop	400
16.6	Soft Magnetic Materials	402
16.7	Hard Magnetic Materials	405
<i>Summary</i>		407
<i>Problems</i>		407
<i>Multiple Choice Questions</i>		409
<i>Sources for Experimental Data</i>		411
<i>Suggestions for Further Reading</i>		411

17 Dielectric Materials **412–426**

17.1	Polarization	413
17.2	Temperature and Frequency Effects	416
17.3	Electric Breakdown	419
17.4	Ferroelectric Materials	420
<i>Summary</i>		423
<i>Problems</i>		424
<i>Multiple Choice Questions</i>		424
<i>Sources for Experimental Data</i>		426
<i>Suggestions for Further Reading</i>		426

Appendix I—Properties of Elements **427–429**

Appendix II—Properties of Engineering Materials **430**

Index **431–439**

Preface

Since the first edition was published, many new features have been added to this book. They include new chapters and new sections on recent topics such as the oxide superconductors, fabrication techniques used in manufacturing integrated circuits, fullerenes, and fracture mechanism maps. In this edition, the text has been updated and rewritten for greater clarity. The diagrams have been improved and drawn using a computer software.

The author is thankful to Mr. Narendra Babu for his assistance in preparing the diagrams. Thanks are also due to the editorial and production team of the Publishers, PHI Learning, in particular, to Mr. K.C. Devasia, for their assistance.

The author is grateful to his colleagues at IIT Delhi—Dr. R.K. Pandey, Dr. S.K. Gupta, Dr. A.N. Kumar and Dr. R. Prasad—for their valuable suggestions.

V. RAGHAVAN

Preface to the First Edition

In keeping with modern trends, the courses in engineering materials or engineering metals and alloys have been replaced by a course in Materials Science in many institutions in India and abroad. Most of the curricula in metallurgy, ceramics and other materials-oriented disciplines have also incorporated a first general course in materials science. This book is intended for use in such courses as well as by the students of applied sciences. Postgraduates who have had no previous exposure to the subject should also find this book useful.

In deciding the level at which this material is to be covered, it has been assumed that the reader has a background in college level physics, chemistry and mathematics. Though not essential, an elementary knowledge of physical chemistry and thermodynamics would be an added advantage.

A reasonably wide coverage in sufficient depth has been attempted, giving the necessary importance to the physical, mechanical, chemical, electrical and magnetic properties. Consistent with the moderate size of the book, the author has tried to emphasize the properties that are more structure-sensitive. Keeping in view the engineering applications, numerous examples of real materials of technological importance have been discussed.

A number of colleagues of the author and over 1500 students who studied this first course during the last 10 years have contributed significantly in class testing and greatly improving this work. In particular, mention must be made of Drs. E.C. Subbarao, D. Chakravorty, L.A. Shepard, M.F. Merriam, C.V. Seshadri, R.K. Mittal and Shri V.M. Kumar and M.L. Gandhi. The author is grateful to Professor N.M. Swani for his encouragement and interest. Special mention must be made of Professor Morris Cohen for providing the author an opportunity to teach materials science courses at the Massachusetts Institute of Technology.

January 1974

V. RAGHAVAN

SI UNITS

Base Units

Quantity	Unit	Symbol
Length l	metre	m
Mass m	kilogram	kg
Time t	second	s
Electric current I	ampere	A
Temperature T	kelvin	K
*Amount of substance n	mole	mol
Luminous intensity	candela	cd

*The mole is the amount of substance that contains as many elementary entities as there are atoms in 0.012 kg of carbon-12. The elementary entities may be atoms, ions, electrons, other particles or groups of particles.

Supplementary Units

Plane angle	radian	rad
Solid angle	steradian	sr

Derived Units with Special Names

Quantity	Special name	Symbol	Equivalence in	
			other derived units	base units
Frequency	hertz	Hz	—	s^{-1}
Force, weight	newton	N	—	$kg\ m\ s^{-2}$
Stress, strength, pressure	pascal	Pa	$N\ m^{-2}$	$kg\ m^{-1}\ s^{-2}$
Energy, work, quantity of heat	joule	J	$N\ m$	$kg\ m^2\ s^{-2}$
Power	watt	W	$J\ s^{-1}$	$kg\ m^2\ s^{-3}$
Electric charge	coulomb	C	—	$A\ s$
Electric potential	volt	V	$W\ A^{-1}$	$kg\ m^2\ s^{-3}\ A^{-1}$
Resistance	ohm	Ω	$V\ A^{-1}$	$kg\ m^2\ s^{-3}\ A^{-2}$
Capacitance	farad	F	$C\ V^{-1}$	$kg^{-1}\ m^{-2}\ s^4\ A^2$
Magnetic flux	weber	Wb	$V\ s$	$kg\ m^2\ s^{-2}\ A^{-1}$
Magnetic flux density	tesla	T	$Wb\ m^{-2}$	$kg\ s^{-2}\ A^{-1}$
Inductance	henry	H	$Wb\ A^{-1}$	$kg\ m^2\ s^{-2}\ A^{-2}$

Prefix Names of Multiples and Submultiples

<i>Factor by which unit is multiplied</i>	<i>Name</i>	<i>Symbol</i>
10^{12}	tera	T
10^9	giga	G
10^6	mega	M
10^3	kilo	k
10^2	hecto	h
10^1	deka	da
10^{-1}	deci	d
10^{-2}	centi	c
10^{-3}	milli	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p
10^{-15}	femto	f
10^{-18}	atto	a

Note: All multiple prefix symbols (except kilo, hecto and deka) are written in capitals and all submultiple symbols are written in lower case letters.

Have a feel for SI units !

Sir Issac Newton, the Apple and SI units

Length unit: 1 metre ~ distance the apple travelled

Time unit: 1 second ~ time of fall of the apple

Force unit: 1 newton ~ weight of the apple

Conventions to be followed when using SI units

1. When writing unit symbols, do not use full stops, plurals, dots or dashes.
2. No degree symbol for kelvin, write as K and not as °K.
3. No kgm (kilogram mass) or kgf (kilogram force). The mass unit is kilogram (kg) and the force unit is newton (N).
4. No space between the prefix symbol and the unit symbol, e.g., meganewton should be written as MN and not as M N.
5. One space between two symbols for clarity, e.g., metre second should be written as m s and not as ms, which means millisecond.
6. All symbols associated with proper names are written with a capital, e.g., A, K, N, etc. When they are written as a word, the lower case is used throughout: ampere, kelvin and newton.
7. Prefix symbols for multiples and submultiples are preferred in steps of 10^3 . Thus, tera (T), giga (G), mega (M), kilo (k), milli (m), micro (μ), nano (n) and pico (p) are preferred prefix symbols.
8. Attach prefix to numerator and not to denominator, e.g., use MN m^{-2} instead of N mm^{-2} , even though both are identical.
9. Three digits are grouped together on either side of the decimal point, e.g.,

1.256 637 83

60 023

Four digit number need not be so grouped, e.g.,

7386

0.6921

Some non-SI units generally accepted

1. degree (celsius), °C.
2. minutes, hours, days, months and years.
3. Angstrom Å for 10^{-10} m.
4. Electronvolt, eV for energy equal to 1.602×10^{-19} J. (1 electronvolt is the kinetic energy acquired by an electron when falling through a potential of 1 volt.)

PHYSICAL CONSTANTS

Avogadro's number	$N = 6.023 \times 10^{23} \text{ mol}^{-1}$
Boltzmann's constant	$k = 1.380 \times 10^{-23} \text{ J K}^{-1}$ $= 8.614 \times 10^{-5} \text{ eV K}^{-1}$
Gas constant	$R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$
Planck's constant	$h = 6.626 \times 10^{-34} \text{ J s}$
Electronic charge	$e = 1.602 \times 10^{-19} \text{ C}$
Electron rest mass	$m_0 = 9.109 \times 10^{-31} \text{ kg}$
Velocity of light	$c = 2.998 \times 10^8 \text{ m s}^{-1}$
Bohr magneton (magnetic moment)	$\mu_B = 9.273 \times 10^{-24} \text{ A m}^2$
Permittivity of free space	$\epsilon_0 = 8.854 \times 10^{-12} \text{ F m}^{-1}$
Permeability of free space	$\mu_0 = 4\pi \times 10^{-7} \text{ H m}^{-1}$ $= 1.257 \times 10^{-6} \text{ H m}^{-1}$
Faraday's constant	$F = 96.49 \text{ kC mol}^{-1} \text{ (of electrons)}$
Atomic mass unit (amu)	$1/(10^3 N) = 1.660 \times 10^{-27} \text{ kg}$
Acceleration due to gravity	$g = 9.81 \text{ m s}^{-2}$

CONVERSION FACTORS

1 inch	=	25.4 mm
1 nm	=	10^{-9} m
1 Å	=	10^{-10} m = 0.1 nm
1°	=	1/57.3 rad
T °C	=	$(T + 273.15)$ K
T °F	=	$5/9(T + 459.67)$ K
1 per °F	=	$9/5$ K ⁻¹
1 kgf	=	9.81 N
1 lb	=	4.45 N
1 dyne	=	10^{-5} N
1 dyne/cm	=	10^{-3} N m ⁻¹
1 atmosphere	=	0.101 325 MN m ⁻²
1 bar	=	10^{-1} MPa
1 psi	=	6.89 kN m ⁻²
1 ksi (10 ³ psi)	=	6.89 MN m ⁻²
1 ton/sq.in.	=	15.46 MN m ⁻²
1 kgf/cm ²	=	98.1 kN m ⁻²
1 kgf/mm ²	=	9.81 MN m ⁻²
1 dyne/cm ²	=	0.1 N m ⁻²
1 torr (mm of Hg)	=	133.3 N m ⁻²
1 kgf/mm ^{3/2}	=	0.310 MN m ^{-3/2}
1 ksi √in	=	1.10 MN m ^{-3/2}
1 eV	=	1.602×10^{-19} J
1 erg	=	10^{-7} J
1 calorie	=	4.18 J
1 eV/entity	=	96.49 kJ mol ⁻¹
1 erg/cm	=	10^{-5} J m ⁻¹
1 erg/cm ²	=	10^{-3} J m ⁻²
1 erg/cm ³	=	0.1 J m ⁻³
1 lb-in/in ²	=	175 J m ⁻²
1 lb/cu.in	=	27 680 kg m ⁻³
1 cm/cm ³	=	10^4 m m ⁻³
1 mole/cm ² /sec	=	10^4 mol m ⁻² s ⁻¹
1 mole/cm ³	=	10^6 mol m ⁻³
1 cm ² /sec	=	10^{-4} m ² s ⁻¹

1 poise	=	0.1 Pa s
1 debye	=	3.33×10^{-30} C m
1 mA/cm ²	=	10 A m ⁻²
1 A hr	=	3.6 kC
1 ohm cm	=	10 ⁻² ohm m
1 mho/cm	=	10 ² ohm ⁻¹ m ⁻¹
1 volt/mil	=	39 370 V m ⁻¹
1 cm ² /volt sec	=	10 ⁻⁴ m ² V ⁻¹ s ⁻¹
1 gauss	=	10 ⁻⁴ Wb m ⁻²
1 oersted	=	79.6 A m ⁻¹

GREEK ALPHABETS

Name	Forms	Sound	Name	Forms	Sound
alpha	α	a	nu	ν	n
beta	β	b	xi	ξ, Ξ	x
gamma	γ, Γ	g	omicron	o	o
delta	δ, Δ	d	pi	π, Π	p
epsilon	ϵ	e	rho	ρ	r
zeta	ζ	z	sigma	σ, Σ	s
eta	η	e	tau	τ	t
theta	θ, Θ	th	upsilon	υ	u
iota	ι	i	phi	ϕ, Φ	ph
kappa	κ	k	khi	χ	kh
lambda	λ, Λ	l	psi	ψ, Ψ	ps
mu	μ	m	omega	ω, Ω	o

In this introductory chapter, we briefly discuss the nature of Materials Science and Engineering. After defining or explaining what Materials Science and Engineering is, we classify engineering materials according to their nature and the various categories of applications. Then we discuss the different levels of the internal structure of materials. Finally, we emphasize the importance of the structure–property relationships in materials, outlining the general approach of the ensuing chapters.

1.1 Materials Science and Engineering

The phrase ‘Materials Science and Engineering’ needs some elucidation for the beginner. The word ‘materials’ here does not refer to all matter in the Universe. If this were so, it would include all the physical sciences and the life sciences—from astronomy to zoology! By including the word ‘inanimate’ in the definition, we can exclude the life sciences from our purview. Further, we can restrict the definition only to matter *useful* to mankind. Even here, the range is too broad for the purposes of the *engineer*. For example, we can list a large number of things useful to man, such as food, medicines, explosives, chemicals, water, steel, plastics and concrete, only some of which qualify as engineering materials. We then have to be more specific and define materials as that part of inanimate matter, which is useful to the engineer in the *practice of his profession*. In the currently understood sense of the term, materials refer only to *solid materials*, even though it is possible to quote a number of examples of liquid and gaseous materials such as sulphuric acid and steam, which are useful to the engineer.

The word ‘science’ in the phrase refers to the physical sciences, in particular to *chemistry* and *physics*. As we confine ourselves mainly to solids in materials science, the subject is related to solid state chemistry and solid state physics. The word ‘engineering’ indicates that the engineering usefulness of the matter under study is always kept in mind, irrespective of whether the basic laws of science can be applied rigorously or not. Where the basic laws cannot be applied, the materials engineer does not give up what is important to him from a *practical point of view*. He uses the best possible approximation, develops

empirical rules, and extrapolates available information to unknown situations. In this respect, materials science and engineering draws heavily from the *engineering sciences such as metallurgy, ceramics and polymer science*. These, in their own time, have grown out of their interaction with the basic sciences of chemistry and physics.

1.2 Classification of Engineering Materials

Having defined the limits of materials that come under our purview, we can classify them in three broad groups according to their nature:

- (i) Metals and alloys
- (ii) Ceramics and glasses
- (iii) Organic polymers.

Metals are familiar objects with a characteristic appearance; they are capable of changing their shape permanently, and have good thermal and electrical conductivity. An *alloy* is a combination of more than one metal. *Ceramics and glasses* are nonmetallic inorganic substances, which are brittle and have good thermal and electrical insulating properties. *Organic polymers* are relatively inert and light, and generally have a high degree of plasticity. Figure 1.1 lists typical examples from each of these three groups of materials. In addition, examples of materials which lie between two groups are also shown.

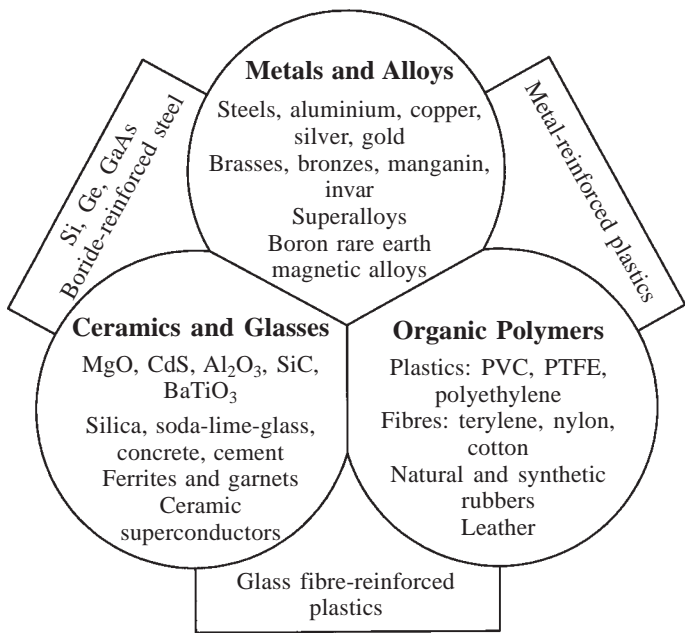


Fig. 1.1 The three major groups of engineering materials.

An alternative way of classifying materials is according to the three major areas in which they are used:

- (i) Structures
- (ii) Machines
- (iii) Devices.

Structures (not to be confused with the internal structure of a material) refer to the objects without moving parts erected by engineers, such as a concrete dam, a steel melting furnace, a suspension bridge and an oil refinery tower. *Machines* include lathes, steam and gas turbines, engines, electric motors and generators. *Devices* are the most recent addition to engineering materials and refer to such innovations as a transistor, a photoelectric cell, piezoelectric pressure gauges, ceramic magnets and lasers.

Invariably, in each category of applications, we find materials from all the three groups described above. To give some examples, an aircraft structure is built of aluminium alloys and plastics; a steel melting furnace is built of refractory oxides and structural steel; safety helmets are made of glass-reinforced plastics. Similarly, we have metal-oxide semiconductors. The block diagram in Fig. 1.2 depicts this interplay between material groups and categories of applications.

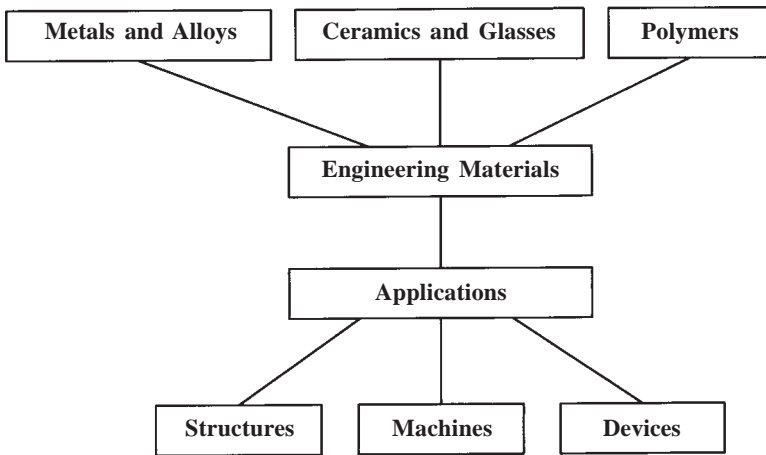


Fig. 1.2 Each category of engineering application requires materials from any or all of the three groups of materials.

1.3 Levels of Structure

The *internal structure* of a material, simply called the structure, can be studied at various levels of observation. The *magnification and resolution* of the physical aid used are a measure of the level of observation. The higher the magnification, the finer is the level. The details that are disclosed at a certain level of observation are generally different from the details disclosed at some other level. Henry Sorby was one of the first to realize this, when he wrote in 1886:

Though I had studied the microscopical structure of iron and steel for many years, it was not until last autumn that I employed what may be called high powers. This was partly because I did not see how this could be satisfactorily done and partly because it seemed to me unnecessary. I had found that in almost every case a power (magnification) of 50 linear showed on a smaller scale as much as one of 200, and this led me to conclude that I had seen the ultimate structure. Now that the results are known, it is easy to see that my reasoning was false, since a power of 650 linear enables us to see a structure of an almost entirely new order.

We have now come a long way since Sorby's time. Magnifications with matching resolutions of a million times linear are now common.

Depending on the level, we can classify the structure of materials as:

Macrostructure

Microstructure

Substructure

Crystal structure

Electronic structure

Nuclear structure.

Macrostructure of a material is examined with naked eye or under a low magnification. The internal symmetry of the atomic arrangements in a crystalline material may reflect in the external form of a crystal such as quartz. Large individual crystals of a crystalline material may be visible to the naked eye, as in a brass doorknob by the constant polishing and etching action of the human hand and sweat.

Microstructure generally refers to the structure as observed under the optical microscope, see Fig. 1.3. This microscope can magnify a structure up to about 1500 times linear, without loss of resolution of details of the structure. The limit

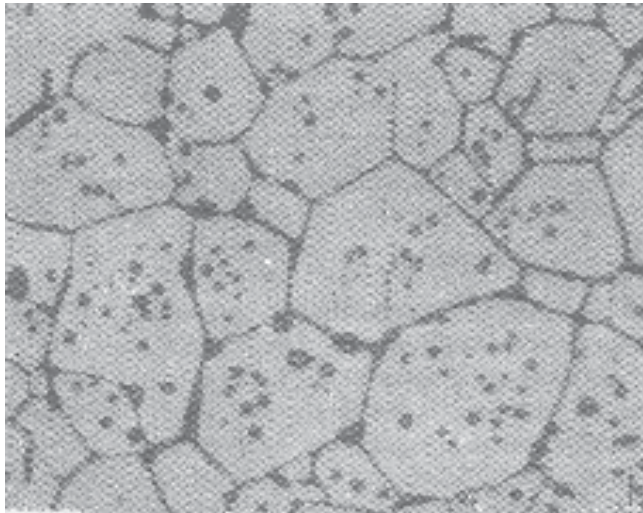


Fig. 1.3 Crystal boundaries in nickel ferrite, Fe_2NiO_4 , magnified 900 times linear.
(W.D. Kingery, *Introduction to Ceramics*, with permission from John Wiley, New York)

of resolution of the human eye is about 0.1 mm (10^{-4} m), that is, the eye can distinguish two lines as separate lines, only when their distance of separation is more than 0.1 mm. The optical microscope can resolve details down to a limit of about 0.1 μm (10^{-7} m).

Substructure refers to the structure obtained by using a microscope with a much higher magnification and resolution than the optical microscope. In an electron microscope, a magnification of 1 000 000 times linear is possible. By virtue of the smaller wavelength of electrons as compared to visible light, the resolving power also increases correspondingly so that much finer details show up in the electron microscope. We can obtain a wealth of additional information on very fine particles or on crystal imperfections such as dislocations. Figure 1.4



Fig. 1.4 Substructure of a Ni-Fe-Cr alloy showing curved dislocation lines, magnified 30 000 times in an electron microscope.

(R.B. Nicholson, *Strengthening Methods in Crystals*, with permission from Applied Science Publishers Ltd., Barking, UK)

shows dislocations imaged in a *transmission electron microscope*. Here, electrons pass through a thin foil of the specimen and the associated diffraction effects produce the image. The electron diffraction patterns obtained along with the photograph of the substructure greatly aid in understanding the processes taking place in materials on such a minute scale.

In a *scanning electron microscope*, an electron beam is scanned across the surface of the specimen, which must be conducting. Back scattered electrons at

each point are collected and displayed at the same scanning rate on a cathode ray tube. The result is an image, much like a television image, of the surface features of the specimen. This image has a very great depth of field so that even a fractured surface without any polishing can be imaged (refer Fig. 12.1e). Magnifications range from $10\times$ to $50\,000\times$.

In the *electron probe microanalyzer*, a beam of electrons bombards the specimen surface. X-rays that have wavelengths characteristic of the elements in the specimen are emitted. Using suitable standards, elements present in the specimen can be identified and their concentration determined from the X-ray intensity. Thus this technique allows microchemical analysis on spots as small as $1\,\mu\text{m}$.

Another modern microscope is the *field ion microscope*. It produces images of individual atoms (Fig. 1.5) and imperfections in atomic arrangements.

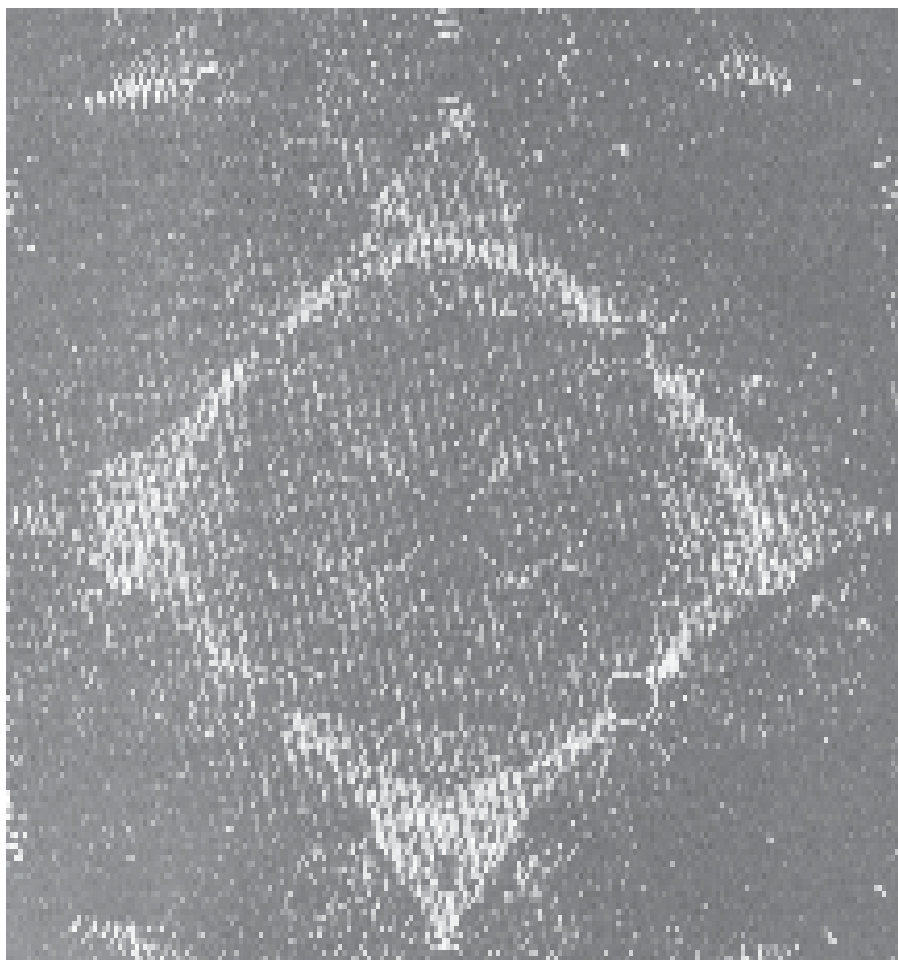


Fig. 1.5 Field-ion micrograph of a hemispherical tip of platinum. The white dots arranged in circles are images of individual atoms.

(Courtesy: E.W. Mueller)

Crystal structure tells us the details of the atomic arrangement within a crystal. It is usually sufficient to describe the arrangement of a few atoms within what is called a unit cell. The crystal consists of a very large number of unit cells forming regularly repeating patterns in space. The main technique employed for determining the crystal structure is the X-ray diffraction.

The *electronic structure* of a solid usually refers to the electrons in the outermost orbitals of individual atoms that constitute the solid. Spectroscopic techniques are very useful in determining the electronic structure.

Nuclear structure is studied by nuclear spectroscopic techniques such as nuclear magnetic resonance (NMR) and Mossbauer studies.

1.4 Structure–Property Relationships in Materials

Until recently, it has been the practice in a course on engineering materials to list the composition, treatment, properties and uses of as many materials as possible. The number and variety of engineering materials and applications have increased tremendously in recent years. Now we have more than a thousand types of steel alone, each with a specific composition, thermal and mechanical history. Therefore, it is impossible to describe an adequate number of engineering materials in one course. Moreover, our knowledge of the internal structure of materials and how this *structure correlates with the properties* has rapidly advanced in recent decades. So, it is more interesting and appropriate to study some of the key factors that determine the structure–property relationships, rather than go for a fully descriptive account of a large number of materials. This approach is adopted in this book. The discussion of a structure-dependent property is usually followed by typical applications.

The levels of structure which are of the greatest interest in materials science and engineering are the microstructure, the substructure and the crystal structure. The chemical, mechanical, electrical and magnetic properties are among the most important engineering properties. We first develop the basic concepts pertaining to the levels of structure. These include concepts in equilibrium and kinetics, the geometry of crystals, the arrangement of atoms in the unit cell of crystalline materials, the substructural imperfections in crystals, and the microstructure of single phase and multi-phase materials. We then discuss how changes in the structure are brought about and how they can be controlled to the best possible advantage. Solid state diffusion and control of phase transformations by heat treatment are the main topics here. In the latter half of the book, corrosion among chemical properties, elastic and plastic deformation among mechanical properties and several electrical and magnetic properties are discussed with numerous examples of typical engineering materials.

The gross composition of a material is important in determining its structure. Yet, for a given gross composition, radical changes in the structure and properties can be brought about by subtle changes in the concentration and

distribution of minute quantities of impurities. The same may also be possible by a thermal or a mechanical treatment that involves no change in the overall composition of the material. Materials Science and Engineering deals more with this kind of changes rather than with the effect of gross composition on the properties.

Suggestions for Further Reading

- A. Street and W. Alexander, *Metals in the Service of Man*, Penguin Books (1976).
- D.L. Weaire and C.G. Windsor (Eds.) *Solid State Science—Past, present and predicted*, Adam Hilger, Bristol, UK (1987).

CHAPTER

2 Equilibrium and Kinetics

In this chapter, we introduce the concept of stability and metastability using the mechanical analog of a tilting rectangular block. We discuss the importance of the indefinite existence of materials in the metastable state. After defining basic thermodynamic functions, we also discuss the statistical nature of entropy and the role of the high-energy fraction of atoms in the statistical distribution in surmounting activation barriers for reactions in materials.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Temperature T	kelvin	K	°C, °F
Pressure P	megapascal or meganewton per square metre	MPa or MN m ⁻²	atmosphere, psi, kg/cm ² , dyne/cm ² , mm of Hg
Internal energy E	joule per mole	J mol ⁻¹	cal/mole, cal/gm
External energy PV			
Enthalpy H			
Gibbs free energy G			
Thermal energy RT			
Activation energy Q	joule per mole per kelvin	J mol ⁻¹ K ⁻¹	cal/mole/°C, cal/gm/°C
Entropy S			
Specific heat C_v, C_p			

Constants

Avogadro's number N	=	$6.023 \times 10^{23} \text{ mol}^{-1}$
Boltzmann's constant k	=	$1.380 \times 10^{-23} \text{ J K}^{-1}$
Gas constant R	=	$8.314 \text{ J mol}^{-1} \text{ K}^{-1}$

2.1 Stability and Metastability

The concept of stability is easily understood by considering a mechanical analog. In Fig. 2.1, a rectangular block of square cross-section is shown in

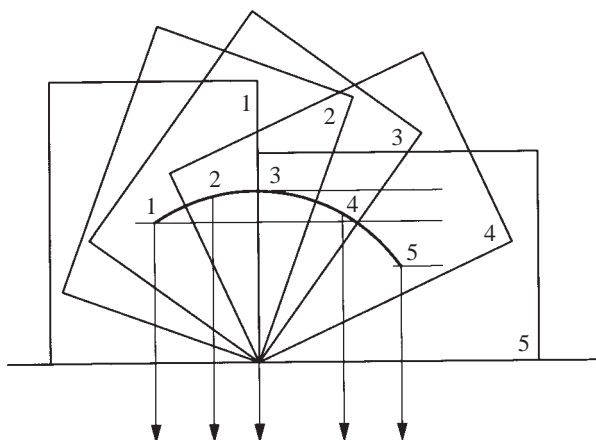


Fig. 2.1 Various positions of a tilting rectangular block illustrate the concept of stability and metastability.

various tilted positions. In position 1, the block is resting on the square base; the arrow from the centre of mass indicates the line along which the weight acts. In position 2, the block is tilted slightly to the right, about one of its edges such that the line along which the weight acts is still within the square base. The centre of mass has moved up due to the tilt. In position 3, the tilt is increased to such an extent that the line of force just falls on the periphery of the base. The centre of mass is now at the maximum possible height from the base. Further tilting lowers it. The line of force now falls outside the square base but within the rectangular base, position 4. On coming to rest on the rectangular face, position 5, the centre of mass is at the lowest possible position for all configurations of the block.

The centres of mass for the various positions of the block are joined by a curve. The potential energy of the block is measured by the height of the centre of the mass from the base. Position 5 corresponds to the lowest potential energy for all configurations and is correspondingly described as the *most stable state* or simply the *stable state*. A system always tends to go towards the most stable state. Position 3 has the maximum potential energy and is called an *unstable state*. Positions 2 and 4 are also unstable states but do not have the maximum energy. Position 1 is called a *metastable state*.

We can use this mechanical analogy to illustrate various equilibrium configurations of a system. Figure 2.2 shows the potential energy of a system as a function of configuration. The potential energy curve has two valleys and a peak. At these positions, the curve has zero slope, that is, the energy does not

vary as a function of configuration for infinitesimally small perturbations. Such configurations are called *equilibrium configurations*. Corresponding to the terminology used for the tilting block, we have stable equilibrium, unstable equilibrium and metastable equilibrium, see Fig. 2.2. Even though the potential energy is a minimum in the metastable state, it is not the lowest for all configurations of the system. Due to the valley position, after small perturbations, the original configuration is restored in both stable and metastable equilibrium. Such restoration does not occur in the case of unstable equilibrium.

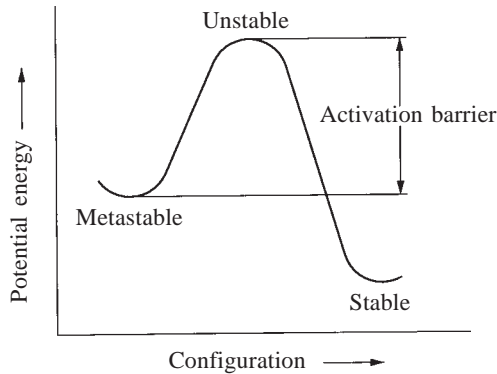


Fig. 2.2 Potential energy as a function of the configuration coordinate.

A metastable state of existence is very common in materials. For example, most metals at room temperature are stable only in the form of an oxide. Oxygen is easily available in the surrounding air. Yet a metal may not combine with oxygen at room temperature (except for a very thin film on the surface). It may exist in the metastable metallic state for an indefinite period. This period could be centuries, as is borne out by the unchanging state of some ancient metallic statues and pillars. This fortunate set of circumstances enables us to use metals in many engineering applications.

In the mechanical analogy we used, the block can be tilted and brought to the most stable configuration starting from a metastable state by an external supply of energy in the form of a hand push or a jiggling base. In materials, the most common source of such energy is the *thermal energy*. *As the temperature of a solid is increased from 0 K, the atoms in the solid vibrate about their mean positions with increasing amplitude.* When the temperature is sufficiently high, rotation of atoms (or small groups of atoms) also becomes possible in some solids. The *translational motion* of atoms past one another is more characteristic of liquids and gases. The energy associated with the vibrations, rotations and translations aids in taking a material from a metastable state to a stable state. In the above example of the oxidation of metals, at sufficiently high temperatures, thermal energy will aid the chemical combination of the metal with the surrounding oxygen, taking it from the metastable to the stable state. This, of course, poses the problem of protecting metals against oxidation, when used at high temperatures.

2.2 Basic Thermodynamic Functions

The concepts of equilibrium and kinetics are intimately associated with the basic thermodynamic parameters. *Pressure* P and *temperature* T are familiar intensive parameters. As opposed to these, we have extensive (or capacity) parameters, which depend on the extent or quantity of the material that comprises the system. The extensive thermodynamic functions are now described.

Internal energy E (also denoted by U) at temperature T is given by

$$E = E_0 + \int_0^T C_v dT \quad (2.1)$$

where E_0 is the internal energy of the material at 0 K and C_v is the specific heat at constant volume. The enthalpy or the heat content of a material H is defined in a similar manner:

$$H = H_0 + \int_0^T C_p dT \quad (2.2)$$

where H_0 is the enthalpy at 0 K and C_p is the specific heat at constant pressure. E and H are related through P and V , where V is the volume of the material:

$$H = E + PV \quad (2.3)$$

PV represents the external energy as opposed to the internal energy and is equal to the work done by the material at constant pressure in creating a volume V for itself. For condensed systems like the liquid and the solid state, at atmospheric pressure, the PV term is negligible so that $E \approx H$. This approximation can be used in most of the problems concerning solid materials.

H_0 represents the energy released when the individual atoms of the material are brought together from the gaseous state to form a solid at 0 K. The gaseous state (where there is no interaction between the atoms) is taken as the reference zero energy state. To indicate that the system has lost energy, H_0 is written with a negative sign. As the temperature increases from 0 K, the material absorbs heat from the surroundings and H increases. The solid melts on reaching the melting point and a further quantity of heat ΔH called the *enthalpy of fusion* is added at the melting temperature. When all the solid has melted, the temperature of the liquid may further increase with the absorption of more energy.

All the energy that a system possesses is not available as work during a chemical change. That part of the energy which can become available as work is called the *Gibbs free energy* (or simply the Gibbs energy). The part which cannot be released as work is called the *bound energy*. Another thermodynamic function called *entropy* defines the relationship between the total energy and the Gibbs energy. At constant pressure, the entropy S of a system is given by

$$S = \int_0^T C_p dT/T \quad (2.4)$$

The entropy of a material is zero at 0 K, in contrast to the enthalpy and the internal energy terms, which have non-zero negative values at 0 K. The entropy increases with increasing temperature. It is a measure of the thermal disorder introduced in the solid, as it is heated above 0 K. The solid state is characterized by the random vibrations of atoms about their mean positions. In the liquid state, the atoms have more freedom and can also move past one another, resulting in greater disorder. Correspondingly, the entropy increases when a solid melts at constant temperature.

Example 2.1 Calculate the entropy increase when one mole of ice melts into water at 0°C.

Solution The latent heat of fusion of ice, $\Delta H = 6.02 \text{ kJ mol}^{-1}$ (80 cal/gm). This heat is absorbed at the constant temperature T_m (0°C). The entropy increase in the process

$$\begin{aligned}\Delta S &= \frac{\text{heat added}}{\text{temperature}} = \frac{\Delta H}{T_m} \\ &= 6.02 \times 10^3 / 273.15 \\ &= 22.04 \text{ J mol}^{-1} \text{ K}^{-1}\end{aligned}$$

In addition to thermal entropy, a system may also possess *configurational entropy*, which is dependent on the configurations of the system. Following Boltzmann's definition, the configurational entropy can be written as

$$S = k \ln w \quad (2.5)$$

where k is Boltzmann's constant and w is the number of different configurations of equal potential energy in which the system can exist. The next section gives a more detailed discussion of configurational and thermal entropies.

The Gibbs energy G is defined in terms of the enthalpy H and the entropy S :

$$G = H - TS \quad (2.6)$$

As already pointed out, the free energy represents the available part of the energy which can be converted to work. As the temperature increases, H increases, but TS increases more rapidly than H and so G decreases with increasing temperature (see Problem 2.3).

The Gibbs energy is used as a criterion of stability. *The most stable state of a material is that which has the minimum Gibbs energy. For a process to occur spontaneously, the Gibbs energy must decrease during the process.* Changes in thermodynamic quantities are always defined as *the final value minus the initial value*:

$$\Delta G = G_{\text{final}} - G_{\text{initial}} \quad (2.7)$$

Then, for a spontaneous process, the free energy change ΔG during the process must be negative. At a constant temperature and pressure, we can write this condition for a spontaneous change as

$$\Delta G = (\Delta H - T \Delta S) < 0 \quad (2.8)$$

Only if there is no change in the entropy of a system during a process, the enthalpy change ΔH can be used in place of ΔG as a criterion of stability. In the example of the tilting block, where no entropy change occurs during the tilt, we were justified in defining the stable state as a state of lowest potential energy (or enthalpy). In general, however, the entropy change may not be negligible. Several chemical reactions are known to be endothermic, that is, they absorb heat during the reaction making ΔH positive, however, they may occur spontaneously, indicating that ΔG is negative. In such cases, $T \Delta S > \Delta H$.

2.3 The Statistical Nature of Entropy

The entropy of a system has been defined by Eqs. (2.4) and (2.5). The meaning of these equations can be understood with reference to physical processes. Taking the configurational entropy first, consider the arrangement of equal numbers of white and black spheres on 16 sites shown in Fig. 2.3. Four possible configurations of these spheres are shown. In Figs. 2.3(a) and 2.3(b), the arrangement is disordered and random. In Fig. 2.3(c), every white sphere is surrounded by black spheres and every black sphere is surrounded by white spheres. In Fig. 2.3(d), all white spheres are separated from the black spheres. It is easy to calculate the total number w of such distinguishable configurations that we can have with these spheres. For generality, if we call the total number of sites as N and the number of white spheres as n , the number of black spheres is $(N - n)$, and it is easily seen that

$$w = \frac{N!}{(N - n)! n!} \quad (2.9)$$

For the above example, substituting $N = 16$ and $n = 8$, we obtain $w = 12\,870$. Among these, only two arrangements shown in Figs. 2.3(c) and 2.3(d) are perfectly ordered and are thus very unlikely to occur in a random choice. The probability of the system existing in a disordered configuration is almost unity.

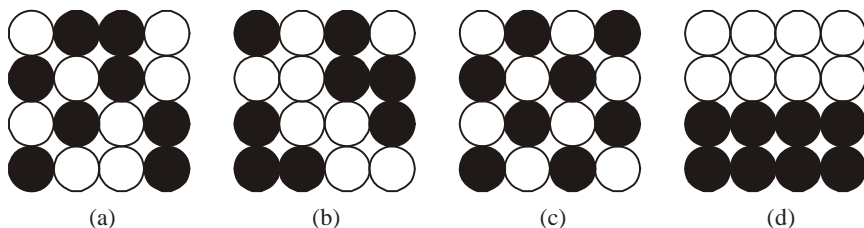


Fig. 2.3 Some possible arrangements of 16 spheres, eight of which are white and the other eight are black.

Equation (2.5) states that the configurational entropy increases as the logarithm of w . One mole of a solid contains more than 10^{23} atoms. If we mix two different kinds of atoms randomly in a solid, we end up with an extremely large number of distinguishable configurations and an appreciable amount of configurational entropy. Since w can never be less than one, the configurational entropy is either zero or positive. It is zero for an absolutely pure solid, consisting of the same kind of atoms on all its sites or for a perfectly ordered solid like a compound.

Consider a mole of atomic sites N in a solid. Let n atoms of B and $(N - n)$ atoms of A be mixed randomly on these sites. The configurational entropy is zero before mixing. The increase in entropy due to mixing is given by

$$\Delta S = S_{\text{final}} - S_{\text{initial}} = k \ln \frac{N!}{(N - n)! n!} \quad (2.10)$$

The following Stirling's approximation is valid for $n \gg 1$:

$$\ln n! = n \ln n - n \quad (2.11)$$

Combining Eqs. (2.10) and (2.11), we obtain

$$\Delta S = k [N \ln N - (N - n) \ln (N - n) - n \ln n] \quad (2.12)$$

Equation (2.4) gives the thermal part of entropy as a function of temperature. At 0 K, we can visualize the atoms of a solid to be at rest, so that there is no disorder due to temperature and, consequently, the entropy is zero. As the temperature increases, the atoms begin to vibrate about their mean positions in the solid with increasing frequency and amplitude. The atoms can be considered to oscillate in three modes corresponding to the three orthogonal directions. For many solids above room temperature, the frequency of these oscillations reaches a constant value, about 10^{13} s^{-1} . In contrast, the amplitude of the oscillations continues to increase with increasing temperature. *The average energy per atom per mode of oscillation is called the thermal energy and is equal to kT* , where T is the temperature in kelvin and k is Boltzmann's constant equal to $1.380 \times 10^{-23} \text{ J K}^{-1}$. For one mole of atoms, the thermal energy becomes

$$NkT = RT \quad (2.13)$$

where N is Avogadro's number and R is the gas constant given by

$$R = Nk = 6.023 \times 10^{23} \times 1.380 \times 10^{-23} = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$$

All the atoms of a solid do not vibrate with the same amplitude and energy at any instant. The vibrating atoms interact with their neighbours and, as a very large number of atoms are involved, the interaction effects are very complex. Consequently, the vibrational energy of any particular atom fluctuates about the average value in a random and irregular way. In other words, there is a statistical distribution of vibrational energies between the atoms. This distribution is the source of the thermal disorder. The entropy given by Eq. (2.4) should be associated with this distribution. The thermal entropy arises from the vibrational energy distribution in a solid, while the configurational entropy arises from the distribution in the configurational arrangements.

In the vibrational energy distribution, the fraction of atoms that possess energy equal to or greater than E is given by the Maxwell-Boltzmann equation:

$$n/N = \exp(-E/kT) \quad (2.14)$$

where n is the number of such atoms and N is the total number of atoms in the solid. The exponential term on the right side, which is a fraction, also gives the probability that any given atom will have energy equal to or greater than E . The Maxwell-Boltzmann equation is basic to the understanding of the equilibrium and kinetic processes in materials.

Example 2.2 Find the fraction of atoms with energy equal to or greater than 1 eV in a solid (i) at room temperature (300 K) and (ii) at 1500 K.

Solution (i) At 300 K, thermal energy $kT = 1.380 \times 10^{-23} \times 300$ J. One electron volt is the kinetic energy acquired by an electron (having a charge of 1.6×10^{-19} coulomb) falling through a potential of 1 V:

$$\begin{aligned} 1 \text{ eV} &= 1.6 \times 10^{-19} \text{ C} \times 1 \text{ V} \\ &= 1.6 \times 10^{-19} \text{ J} \end{aligned}$$

So,
$$kT = 1.38 \times 10^{-23} \times 300 / (1.6 \times 10^{-19}) \text{ eV}$$
$$= 0.026 \text{ eV}$$

$$\begin{aligned} n/N &= \exp(-1/0.026) = \exp(-38.5) \\ &= 10^{-16.70} = 2 \times 10^{-17} \end{aligned}$$

(ii) At 1500 K, $kT = 0.129 \text{ eV}$

$$n/N = \exp(-7.75) = 4.3 \times 10^{-4}$$

With a five-fold increase in temperature, the fraction of the highly energetic atoms has increased by about 13 orders of magnitude.

2.4 The Kinetics of Thermally Activated Processes

Arrhenius first measured the rate of a chemical reaction as a function of temperature. He found that the rate is an exponential function of temperature according to the following equation:

$$\text{Rate} = A \exp\left(-\frac{Q}{RT}\right) \quad (2.15)$$

where A is a pre-exponential constant and Q is called the activation energy. In Fig. 2.4, the logarithm of the rate of a reaction is plotted as a function of the

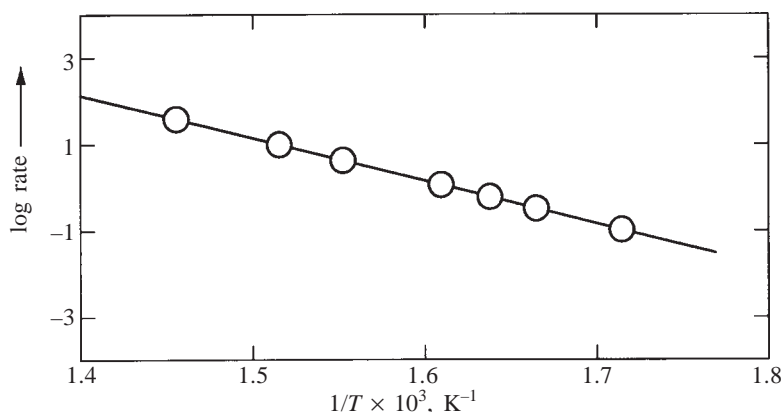
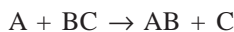


Fig. 2.4 The Arrhenius plot for the reaction: $\text{I}_2 + \text{H}_2 \rightleftharpoons 2\text{HI}$.

reciprocal of temperature. Such a plot is called the *Arrhenius plot*. The straight line relationship is evident. The slope of the straight line is equal to $-Q/R$ and the *intercept* on the y-axis is the logarithm of A . This exponential dependence of a reaction rate on temperature can be justified as follows: Consider a reaction between different atomic species A, B and C. Let atom A interact with molecule BC to produce molecule AB and atom C:



For this reaction to go in the forward direction as indicated, A and BC must come into contact, the bond between B and C must be broken and the bond between A and B must be formed. This usually involves an intermediate step through which the reaction has to proceed:



The *complex species* $(\text{ABC})^*$ forms in the intermediate step. The exact nature of this complex is not a matter of concern to us here. The important point is that this complex has a high energy. The energy variation along the path of the reaction is shown schematically in Fig. 2.5. The *activated complex* is at the peak

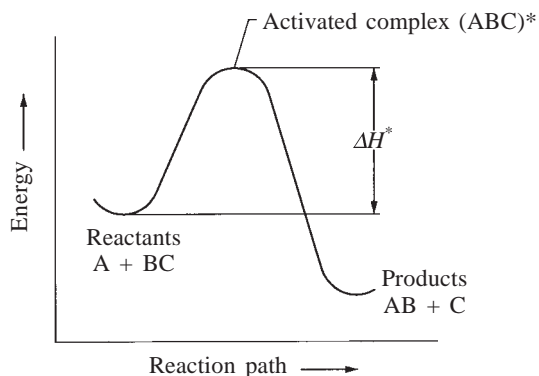


Fig. 2.5 Energy variation along the reaction path. The activation barrier is ΔH^* .

of this energy curve. This figure can be compared with the potential energy variation, when the rectangular block is tilted. The activated complex here corresponds to the unstable position of the tilting block.

For the reaction to occur, the reacting species A and BC must first form the activated complex, that is, they must reach the energy peak. Once the activated state is reached, the reaction can instantly proceed to the final step. Just one out of the 10^{13} vibrations that occur every second can take the reaction down the energy hill to the final product AB and C.

The *average thermal energy* of an atom is insufficient to surmount a typical activation barrier, which has a height of 1 eV. The average thermal energy of any mode reaches 1 eV, only at an extremely high temperature, 12 000 K. Yet many processes in materials that are dependent on thermal energy proceed vigorously at much lower temperatures. This dilemma is resolved by considering the Maxwell-Boltzmann energy distribution. Equation (2.14) tells us that there is always a small fraction of species with energies *much greater than the average*. We then have to rely on this small fraction to surmount the activation barrier. The small fraction of atoms, which has energy equal to or greater than the activation barrier ΔH^* , would reach the final state at the very start, thus making the reaction to proceed in the forward direction by a small amount. The consequent depletion that occurs in the number of the high energy species is replenished immediately, with the restoration of the equilibrium statistical distribution. The reaction thus continues as a function of time. The rate of the reaction is proportional to the number of species n with energy equal to or greater than that of the activation barrier, as well as to the vibrational frequency ν . The vibrational frequency determines the rate of the final step from the activated state to the product species.

$$\text{Rate} \propto n\nu \propto N\nu \exp\left(-\frac{\Delta H^*}{kT}\right) \quad (2.16)$$

where N is the total number of species. The similarity in form of this equation with the Arrhenius equation (2.15) is evident. The height of the barrier ΔH^* is to be equated to the experimental activation energy Q . The vibrational frequency ν and the total number of species N form part of the pre-exponential term A .

SUMMARY

1. The concept of stability and metastability can be understood with reference to various positions of a rectangular block.
2. As we tilt the block from the metastable position to the stable position, we pass through the high-energy unstable state.

3. Pressure and temperature are intensive thermodynamic parameters, as opposed to internal energy, enthalpy, entropy and free energy, whose values depend on the quantity of the material.
4. Entropy defines the relationship between the total energy and the free energy of a material.
5. Entropy can be subdivided into thermal and configurational parts. It measures the thermal disorder, determined by the random distribution of vibrational energies between atoms and the configurational disorder, determined by the variations in configurational arrangements.
6. The Maxwell-Boltzmann energy distribution law is basic to the understanding of the equilibrium and kinetic processes in materials. It provides an explanation for the exponential dependence on temperature of the rate of a thermally activated process.

PROBLEMS

- 2.1** Calculate the increase in the enthalpy and the entropy of copper, as it is heated from room temperature (300 K) to 1000 K. The specific heat in this temperature range is given by

$$C_p = 22.61 + 6.27 \times 10^{-3}T \text{ J mol}^{-1} \text{ K}^{-1}.$$

Answer: $\Delta H = 18.7 \text{ kJ mol}^{-1}$ and $\Delta S = 31.6 \text{ J mol}^{-1} \text{ K}^{-1}$.

- 2.2** Calculate the entropy change and the free energy change during the melting of gold at its melting point. The enthalpy of fusion for gold is 12.6 kJ mol^{-1} .

Answer: $\Delta S = 9.43 \text{ J mol}^{-1} \text{ K}^{-1}$; $\Delta G = 0$.

- 2.3** Show that $\left(\frac{\partial G}{\partial T}\right)_p = -S$.

Make a schematic plot of G as a function of temperature for a solid, starting from 0 K.

- 2.4** Show that the entropy of mixing equal numbers of two different kinds of atoms on one mole of fixed atomic sites is $5.76 \text{ J mol}^{-1} \text{ K}^{-1}$.
- 2.5** Calculate the fraction of atoms with energies equal to or greater than 2 eV at 300 K and 1500 K. Compare these values with those calculated in Example 2.2.

Answer: 4.0×10^{-34} ; 1.9×10^{-7} .

- 2.6** Show that the logarithm of the ratio of the reaction rates at two temperatures T_1 and T_2 is equal to $Q \Delta T / (RT_2 T_1)$. From this, find the temperature increase

starting from room temperature required to double the reaction rate for a value of Q equal to (i) 100 kJ mol^{-1} and (ii) 200 kJ mol^{-1} .

Answer: (i) $\sim 5^\circ\text{C}$, (ii) $\sim 2.5^\circ\text{C}$.

- 2.7** Find the activation energy from an Arrhenius plot for a reaction that requires the following times for completion at the indicated temperatures:

Temperature, K	Time, s
600	9360
700	27.5
800	0.73
900	0.01

Answer: $\sim 200 \text{ kJ mol}^{-1}$.

- 2.8** A reaction takes 500 min at 10°C for completion. It takes 1 min at 80°C . Find the time it would take at 40°C .

Answer: 25 min.

- 2.9** The vibrations of atoms in a solid can be considered as elastic waves. The limiting wavelength of the elastic waves can be taken to be the interatomic spacing, which has a value of about 10^{-10} m . Estimate in order of magnitude the velocity of elastic waves in a solid above room temperature.

Answer: 1 km s^{-1} .

- 2.10** In Eq. (2.12), show that $(d\Delta S/dn) \rightarrow \infty$, as $n \rightarrow 0$. Explain why it is very difficult to remove the last traces of impurities during purification of a material.

- 2.11** A reaction with activation energy equal to 100 kJ mol^{-1} takes 50 min for completion at 300 K. At what temperature will it be complete in 5 min?

Answer: 318.3 K.

MULTIPLE CHOICE QUESTIONS

- The number of ways of arranging 6 Ni atoms and 6 Cu atoms on 12 atomic sites are
A. 432 B. 924 C. 12870 D. 36
- The number of ways of arranging 7 white spheres and 7 black spheres on 15 sites are
A. 3432 B. 12870 C. 51480 D. none of these
- The entropy of mixing of 0.4 mole of Au atoms, 0.4 mole of Ag atoms and 0.2 mole of Cu atoms on one mole of sites in $\text{J mol}^{-1} \text{ K}^{-1}$ is
A. 8.77 B. 5.76 C. 5.59 D. 4.16

4. The entropy of mixing of 0.5 mole of Ni atoms and 0.49 mole of Cu atoms on 1 mole of sites in $\text{J mol}^{-1} \text{K}^{-1}$ is
A. 5.76 B. 5.79 C. 5.85 D. 6.17
5. The slope of the Gibbs energy G versus T curve at 0 K and constant pressure is
A. 0 B. H_0 C. S_0 D. $-H$
6. The thermal energy at room temperature per mode of oscillation of one mole of atoms is
A. 2.49 kJ B. 7.48 kJ C. 1.24×10^{-20} J D. 4.14×10^{-21} J
7. The free energy change during melting of ice at 0°C is equal to
A. enthalpy of melting–entropy of melting
B. 0 C. 273
D. can't say without more data
8. The entropy becomes zero at 0°C for a
A. pure element B. perfect crystal
C. random solid solution D. none of these
9. The number of atoms in one mole of an elemental crystal possessing energy equal to or greater than 1 eV at 1000°C is
A. 0 B. 1.1×10^{-4} C. 6.67×10^{19} D. 6.02×10^{23}
10. A reaction takes 500 min at 10°C and 25 min at 40°C for completion. The activation energy for the reaction in kJ/mol is
A. 73.5 B. 33.4 C. 96.49 D. 100
11. A reaction takes 500 and 1 min respectively at 10° and 80°C . The time it would take at 50°C is
A. 25 min B. 15 min C. 10 min D. 6 min
12. The rate of a thermally activated reaction is 5 at 5°C and 10 at 10°C . Its rate at 20°C in the same units is
A. 14.1 B. 20 C. 37.3 D. 40
13. The temperature increase from 25°C required to triple a reaction rate is 10°C . The activation energy (kJ mol^{-1}) for the reaction is
A. 0.8 B. 53 C. 106 D. 84

Answers

- | | | | | |
|-------|-------|-------|------|-------|
| 1. B | 2. C | 3. A | 4. D | 5. A |
| 6. A | 7. B | 8. D | 9. C | 10. A |
| 11. C | 12. C | 13. D | | |

Sources for Experimental Data

- I. Barin, O. Knacke, and O. Kubaschewski, *Thermochemical Properties of Inorganic Substances*, Springer-Verlag, Berlin (1973) and Supplement (1977).
- R. Hultgren et al. (Eds.), *Selected Values of the Thermodynamic Properties of the Elements*, American Society for Metals, Metals Park, Ohio (1973).

Suggestions for Further Reading

- E.A. Guggenheim, *Thermodynamics*, North-Holland, Amsterdam (1967).
- C.H.P. Lupis, *Chemical Thermodynamics of Materials*, North-Holland, Amsterdam (1983).

Materials can be broadly classified as crystalline and noncrystalline solids. In a crystal, the arrangement of atoms is in a periodically repeating pattern, whereas no such regularity of arrangement is found in a noncrystalline material. A crystalline solid can be either a single crystal, where the entire solid consists of only one crystal, or an aggregate of many crystals separated by well-defined boundaries. In the latter form, the solid is said to be *polycrystalline*.

In this chapter, we introduce the elementary concepts of crystal geometry. We first discuss the geometry of arranging points in regular patterns in space. A crystal can be generated starting from such patterns. We then introduce the conventions associated with representing directions and planes in crystals. We describe in some detail the experimental method of determining simple crystal structures using the x-ray diffraction technique.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Lattice parameter a Atomic diameter Interplanar spacing d Wavelength of radiation λ	nanometre	nm	Å
Camera radius R	millimetre	mm	inch
Bragg angle θ	radian	rad	°(degree)

Note: Lengths of atomic dimensions can be expressed in units of nanometre. However, another submultiple of the metre, the angstrom unit (Å), is still used extensively, even though it is not a preferred submultiple. We will continue to use the angstrom unit, as it is convenient and as the conversion to metre is straightforward. Likewise, we will continue to use °(degree) for angle.

GEOMETRY OF CRYSTALS

3.1 The Space Lattices

Before discussing the periodic patterns of atomic arrangements in crystals, we need to look into arrangements of *points* in space in periodically repeating patterns. This leads us to the concept of a space lattice. A space lattice provides the framework with reference to which a crystal structure can be described.

A *space lattice* is defined as an infinite array of points in three dimensions in which every point has surroundings *identical* to that of every other point in the array. As an example, for ease of representation on paper, consider a two-dimensional square array of points shown in Fig. 3.1. By repeated translation of

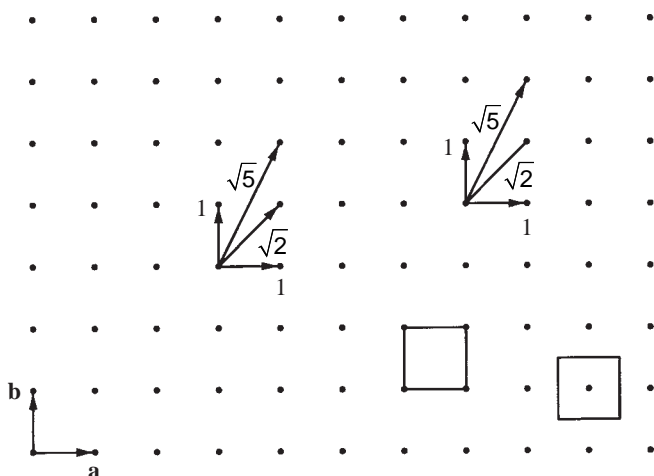


Fig. 3.1 A two-dimensional square array of points gives a square lattice. Two ways of choosing a unit cell are illustrated.

the two vectors **a** and **b** on the plane of the paper, we can generate the square array. The magnitudes of **a** and **b** are equal and can be taken to be unity. The angle between them is 90° ; **a** and **b** are called the *fundamental translation vectors* that generate the square array. To ignore end effects near the boundary, we will assume that the array can be extended infinitely. If we locate ourselves at any point in the array and look out in a particular direction that lies on the plane of the paper, the scenery is the same, irrespective of where we are. Consider the immediate surroundings of a point in the array. If we look due north or due east from this point, we see another point at a distance of 1 unit. Along northeast, we see the nearest point at a distance of $\sqrt{2}$ units and along north-northeast, the nearest point is at a distance of $\sqrt{5}$ units. As this is true of every point in the array, the array satisfies the definition given above and can be called a two-dimensional square lattice.

Example 3.1 Draw a two-dimensional pentagonal lattice.

Solution A regular pentagon has an interior angle of 108° . As 360° is not an integral multiple of 108° , pentagons cannot be made to meet at a point bearing a constant angle to one another. Hence, a pentagonal lattice is not possible. On the other hand, a square or a hexagonal two-dimensional lattice is possible.

A space lattice can be defined by referring to a unit cell. The unit cell is the smallest unit which, when repeated in space indefinitely, will generate the space lattice. In the above example of the square lattice, the unit cell is the square obtained by joining four neighbouring lattice points, as shown in Fig. 3.1. Since every corner of this square is common to four unit cells meeting at that corner, the effective number of lattice points in the unit cell is only one. Alternatively, the unit cell can be visualized with one lattice point at the centre of the square and with none at the corners (see Fig. 3.1).

A three-dimensional space lattice is generated by repeated translation of three noncoplanar vectors, **a**, **b** and **c**. It so turns out that there are only 14 distinguishable ways of arranging points in three-dimensional space such that each arrangement conforms to the definition of a space lattice. These 14 space lattices are known as *Bravais lattices*, named after their originator. They belong to seven *crystal systems* and are listed in Table 3.1 according to the crystal system.

The *cubic system* is defined by three mutually perpendicular translation vectors **a**, **b**, and **c**, which are equal in magnitude. The angle between **b** and **c** is α , the angle between **c** and **a** is β , and that between **a** and **b** is γ . This notation about angles is general and should be consistently followed. As shown in Table 3.1, there are three space lattices in the cubic crystal system: the simple cubic, the body centred cubic, and the face centred cubic space lattices.

Example 3.2 Derive the effective number of lattice points in the unit cell of the three cubic space lattices.

Solution The unit cell in all these three cases is the cube. The corners of a cube are common to eight adjacent cubes. The faces are common to two adjacent cubes. The body centre is not shared by any other cube. So, the effectiveness of a corner lattice point is $1/8$, that of a face centred lattice point is $1/2$ and that of the body centre is 1. Referring to Table 3.1, we can write:

Space lattice	Abbreviation	Effective number of lattice points in unit cell
Simple cubic	SC	1
Body centred cubic	BCC	2
Face centred cubic	FCC	4

TABLE 3.1
The Bravais Lattices

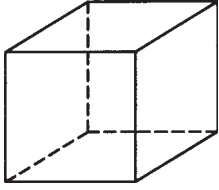
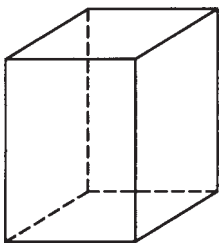
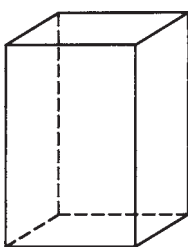
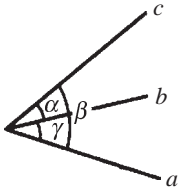
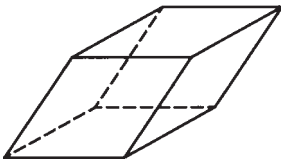
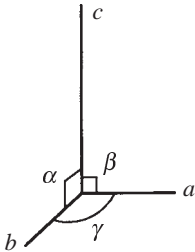
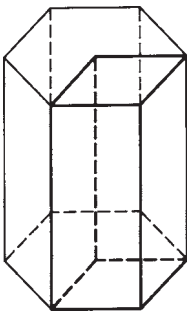
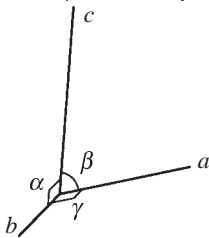
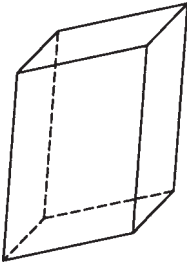
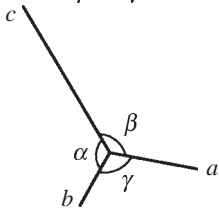
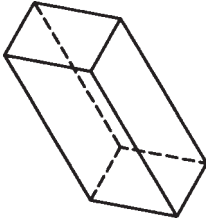
Crystal system	Space lattice	Unit cell
<p>I. Cubic</p> $a = b = c$ $\alpha = \beta = \gamma = 90^\circ$	<p>(1) Simple (Lattice points at the eight corners of the unit cell)</p> <p>(2) Body centred (Points at the eight corners and at the body centre)</p> <p>(3) Face centred (Points at the eight corners and at the six face centres)</p>	
<p>II. Tetragonal</p> $a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	<p>(4) Simple (Points at the eight corners of the unit cell)</p> <p>(5) Body centred (Points at the eight corners and at the body centre)</p>	
<p>III. Orthorhombic</p> $a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	<p>(6) Simple (Points at the eight corners of the unit cell)</p> <p>(7) End centred (Also called side centred or base centred) (Points at the eight corners and at two face centres opposite to each other)</p> <p>(8) Body centred (Points at the eight corners and at the body centre)</p> <p>(9) Face centred (Points at the eight corners and at the six face centres)</p>	

TABLE 3.1 (cont.)
The Bravais Lattices

Crystal system	Space lattice	Unit cell
IV. Rhombohedral $a = b = c$ $\alpha = \beta = \gamma \neq 90^\circ$ 	(10) Simple (Points at the eight corners of the unit cell)	
V. Hexagonal $a = b \neq c$ $\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$ 	(11) Simple [(i) Points at the eight corners of the unit cell outlined by thick lines or (ii) Points at the twelve corners of the hexagonal prism and at the centres of the two hexagonal faces]	
VI. Monoclinic $a \neq b \neq c$ $\alpha = \gamma = 90^\circ \neq \beta$ 	(12) Simple (Points at the eight corners of the unit cell) (13) End centred (Points at the eight corners and at two face centres opposite to each other)	
VII. Triclinic $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90^\circ$ 	(14) Simple (Points at the eight corners of the unit cell)	

Crystals have inherent *symmetry*. A cubic crystal is said to have a four-fold rotation symmetry about an axis passing through the centres of two opposite faces of the unit cube. During each complete rotation about this axis, the crystal passes through identical positions in space four times. The rotational, translational and reflection symmetry operations constitute the symmetry elements of a crystal. The crystal systems in Table 3.1 are arranged in order of decreasing symmetry, the cube being the most symmetric and the triclinic system being the least symmetric. The details of symmetry elements of a crystal are given in books on crystallography and will not be covered here.

After the cubic system, the next less symmetric crystal system is the *tetragonal system*. It is defined by three mutually perpendicular vectors, only two of which are equal in magnitude. There are two space lattices here.

Space lattice	Abbreviation	Effective number of lattice points in unit cell
Simple tetragonal	ST	1
Body centred tetragonal	BCT	2

It is interesting to note that there is no face centred tetragonal space lattice. Any array of lattice points that can be represented by an FCT cell can equally well be described by a BCT cell, as illustrated in Fig. 3.2. When there are two such alternatives of the same crystal system available to describe the same array of lattice points, the unit cell which has the smaller number of lattice points is chosen for the Bravais list.

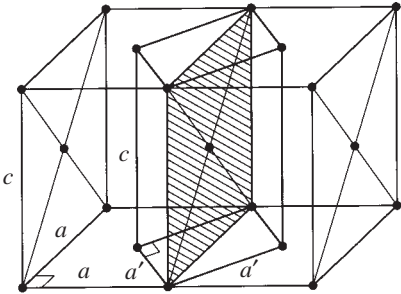


Fig. 3.2 An array of lattice points that fit in an FCT unit cell should be represented by a BCT cell (outlined by thick lines inside two adjacent FCT cells). [For clarity, the face centred atoms on the front and back faces are omitted.]

An orthorhombic cell is defined by three unequal but mutually perpendicular translation vectors. There are four orthorhombic space lattices.

Space lattice	Abbreviation	Effective number of lattice points in unit cell
Simple orthorhombic	SO	1
End centred orthorhombic	ECO	2
Body centred orthorhombic	BCO	2
Face centred orthorhombic	FCO	4

As a final example, let us examine the unit cell of the hexagonal system. In order that the hexagonal symmetry becomes evident, we can take the unit cell to be a regular hexagonal prism, see Table 3.1. The effective number of lattice points in this unit cell is 3. For generating the entire space lattice by translation of the unit cell, a smaller cell with only one lattice point is used, refer Table 3.1.

Example 3.3 List the lattice (unit cell) parameters required to specify fully the unit of each crystal system.

Solution

Crystal system	To be specified		Total number of parameters
	<i>axes</i>	<i>angles</i>	
Cubic	a	—	1
Tetragonal	a, c	—	2
Orthorhombic	a, b, c	—	3
Rhombohedral	a	α	2
Hexagonal	a, c	—	2
Monoclinic	a, b, c	β	4
Triclinic	a, b, c	α, β, γ	6

3.2 Space Lattices and Crystal Structures

A space lattice is combined with a *basis* to generate a *crystal structure*.

$$\text{Space lattice} + \text{Basis} \rightarrow \text{Crystal structure}$$

In many elemental crystals, the basis is simple and consists of one atom per lattice point. In such cases, the crystal is generated by just positioning one atom of the element at each lattice point. For example, the crystal structures of chromium and copper are generated as:

BCC space lattice + 1 Cr atom per lattice point \rightarrow BCC crystal of Cr

FCC space lattice + 1 Cu atom per lattice point \rightarrow FCC crystal of Cu

These crystals are called *monoatomic crystals*, to denote the fact that the basis is one atom per lattice point. This adjective is often omitted. A BCC crystal means a monoatomic BCC crystal, unless otherwise stated.

Figure 3.3 shows unit cells of monoatomic cubic crystals. The atoms are visualized to be hard balls and are shown such that any *two nearest neighbours touch each other*. This is consistent with the definition that *the diameter of an atom in an elemental crystal is the closest distance of approach between two atoms*. In the simple cubic unit cell, this distance is the cube edge. In the BCC cell, it is half of the body diagonal. In the FCC cell, it is half of the face diagonal. If a is the lattice parameter (cube edge), we can write:

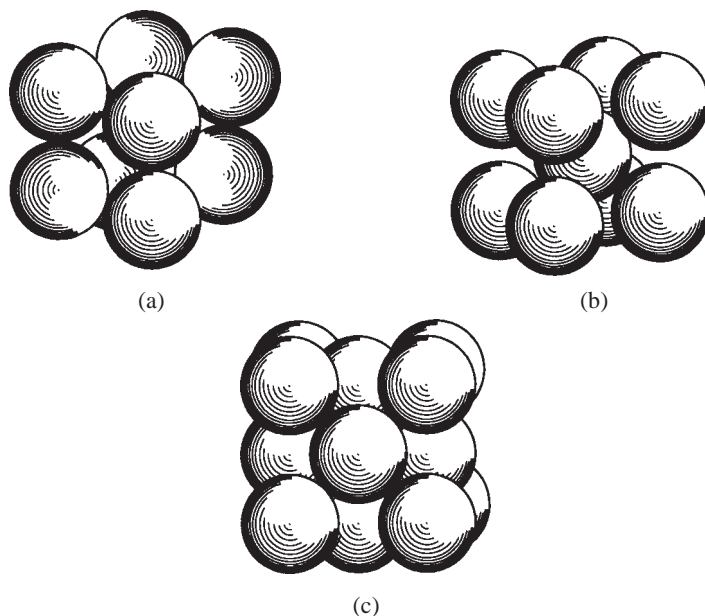


Fig. 3.3 Unit cells of monoatomic crystals of (a) simple cubic (SC), (b) body centred cubic (BCC), and (c) face centred cubic (FCC) structure.

Crystal (monoatomic)	Number of nearest neighbours	Closest distance of approach or atomic diameter
SC	6	a
BCC	8	$a\sqrt{3}/2$
FCC	12	$a/\sqrt{2}$

The student should check the number of nearest neighbours listed above for each case.

In principle, an infinite number of crystal structures can be generated by combining different bases and different lattice parameters with the same space lattice. In Fig. 3.4, three different bases are combined with the simple cubic lattice. In Fig. 3.4a, the crystal is monoatomic, with just one atom at each lattice point. For clarity, neighbouring atoms are shown separately. Figure 3.4b illustrates a *molecular crystal*, with a diatomic molecule at each lattice point. The centre of the larger atom of the molecule coincides with a lattice point, while the smaller atom is not at a lattice point. In molecular crystals, the basis is fully defined by giving the number and types of atoms, the internuclear distance of separation between neighbours in the molecule and the orientation of the molecule in relation to the unit cell. In Fig. 3.4c, the corner atoms of the cube are of one type, but the atom at the body centre is of a different type. The basis is two atoms, the larger one in this case at a lattice point and the smaller one positioned half-way along the body diagonal, at the body centre, which is *not* a

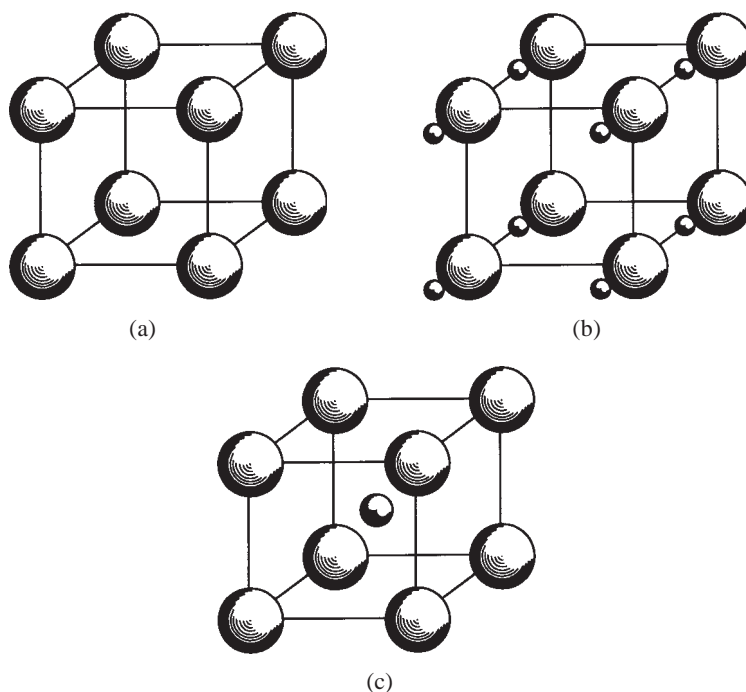


Fig. 3.4 Three different crystal structures referred to the same simple cubic lattice. [For clarity, neighbouring atoms or molecules are shown separated.]

lattice point. In the crystal, of course, the unit cell can be shifted such that the body centre becomes a lattice point and the body corners are no longer lattice points. This crystal should not be confused with the monoatomic BCC crystal, where the body corner and the body centre atoms are of the same type.

The number of crystal structures known to exist runs into thousands. This indicates that there can be more complex bases than those we have considered above. For example, one crystal form of manganese has the structure referred to the BCC space lattice, with 29 atoms grouped together at each lattice point. In protein structures, the number of atoms in the basis may be as high as 10 000! Obviously, the description of such bases would include a number of complex details.

3.3 Crystal Directions and Planes

It is necessary to use some convention to specify directions and planes in a crystal. For this purpose, the system devised by Miller known as Miller indices is widely used. In Fig. 3.5, the vector \mathbf{r} , passing through the origin o to a lattice point, can be expressed in terms of the fundamental translation vectors \mathbf{a} , \mathbf{b} and \mathbf{c} , which form the crystal axes, as

$$\mathbf{r} = r_1\mathbf{a} + r_2\mathbf{b} + r_3\mathbf{c} \quad (3.1)$$

where r_1 , r_2 and r_3 are integers. The c -axis is not shown in the figure as \mathbf{r} is assumed to lie on the ab plane. The components of \mathbf{r} along the three axes are: $r_1 = 2$, $r_2 = 3$ and $r_3 = 0$. Then the *crystal direction* denoted by \mathbf{r} is written as $[230]$ in the Miller notation, with *square brackets* enclosing the indices.

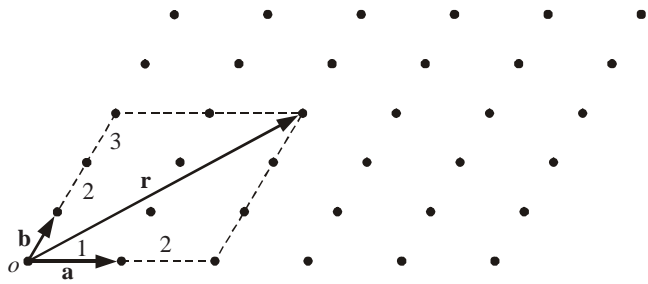


Fig. 3.5 The Miller indices of the crystal direction denoted by vector \mathbf{r} are $[230]$.

If there is a negative component along a crystal axis such as -2 , it is written as $\bar{2}$ and read as bar 2. A family of directions is obtained by all possible combinations of the indices, both positive and negative. The family $[230]$, $[203]$, $[\bar{2}0\bar{3}]$, $[30\bar{2}]$, $[\bar{3}20]$, etc., is represented by $\langle 230 \rangle$, where the *angular brackets* $\langle \rangle$ denote the entire family. Such a convention of representing a family is very convenient for cubic crystals.

Example 3.4 Find the family of crystal directions represented by cube edges, face diagonals and body diagonals of the unit cube. Give the number of members in each family.

Solution

Direction	Miller indices	Number in the family
Cube edge	$\langle 100 \rangle$	6
Face diagonal	$\langle 110 \rangle$	12
Body diagonal	$\langle 111 \rangle$	8

It is left as an exercise to the student to write down the Miller indices of each member of the three families.

The magnitude of the vector \mathbf{r} gives the magnitude of that crystal direction. The crystal directions $[230]$, $[460]$ and $[1 \frac{1}{2} 0]$ all have the same direction, but different magnitudes. Since Miller indices for directions are usually specified as the smallest possible integers, the differences in magnitude for the above three directions are indicated using the following convention:

$$[230], 2[230] \text{ and } \frac{1}{2}[230]$$

The Miller indices of a crystal plane are determined as follows. Referring to the plane shown in Fig. 3.6:

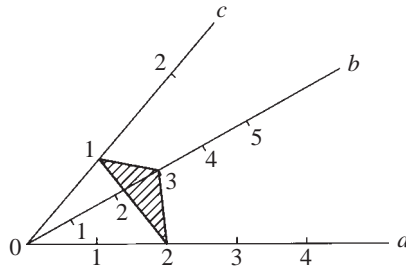


Fig. 3.6 A crystal plane making intercepts 2, 3 and 1 on the crystal axes *a*, *b* and *c*. Its Miller indices are (326).

Step 1	Find the intercepts of the plane along the axes <i>a</i> , <i>b</i> and <i>c</i> (the intercepts are measured as multiples of the fundamental vectors)	2 3 1
Step 2	Take reciprocals of the intercepts [#]	$\frac{1}{2} \frac{1}{3} 1$
Step 3	Convert into smallest integers in the same ratio	3 2 6
Step 4	Enclose in parentheses	(326)

The factor that results in converting the reciprocals to integers may be indicated outside the brackets, but it is usually omitted. The family of planes with members (236), (263), ($\overline{3}62$), (326), ($\overline{6}32$); etc., is denoted by {326}, the curly brackets { } standing for the family.

Example 3.5 Find the Miller indices of the direction *r* and the plane indicated by unit normal *s* in Fig. 3.7.

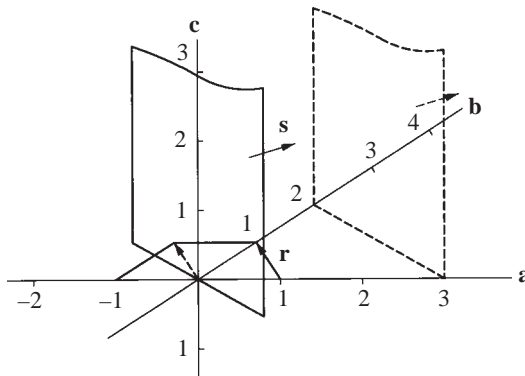


Fig. 3.7 A crystal direction *r* and a crystal plane denoted by unit normal *s*.

[#]The reciprocal procedure avoids the intercept of ∞ for a plane parallel to an axis, by making it 0.

Solution The direction \mathbf{r} does not pass through the origin in the figure. As the choice of origin is arbitrary, shift the origin so that \mathbf{r} passes through it or, alternatively, draw a vector parallel to \mathbf{r} and passing through the origin. Then the Miller indices are found to be $[\bar{1}10]$.

The plane s passes through the origin. Draw a parallel plane that makes the smallest integral intercepts on the coordinate axes. The intercepts for the parallel plane drawn in Fig. 3.7 are 3, 2 and ∞ , so that the Miller indices are given by (230).

Example 3.6 Draw a (110) and a $(\bar{1}\bar{1}1)$ plane inside a cubic unit cell. Determine the Miller indices of the direction that is common to both these planes.

Solution In Fig. 3.8, the plane (110) is sketched using the origin O . Shifting the origin to the opposite corner of the bottom face, the plane $(\bar{1}\bar{1}1)$ is sketched. The direction common to the two planes $[\bar{1}10]$ is also shown.

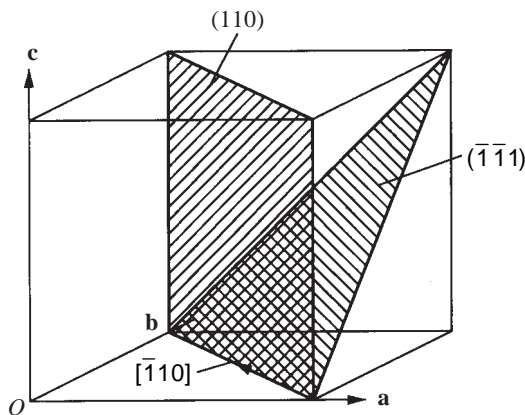


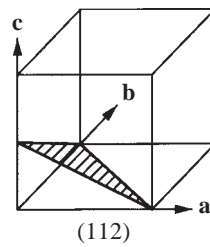
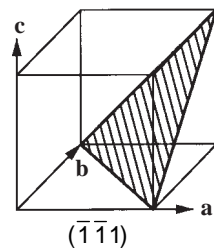
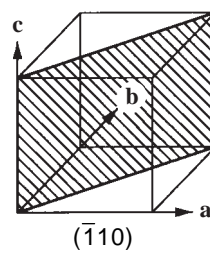
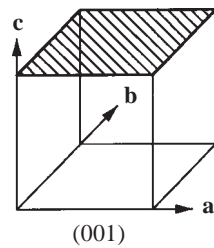
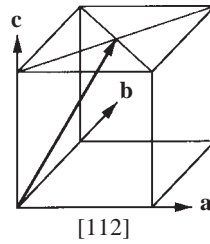
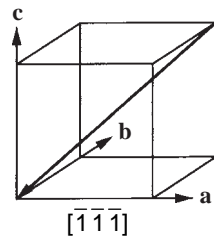
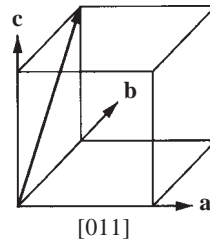
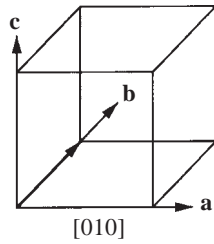
Fig. 3.8 (110) and $(\bar{1}\bar{1}1)$ planes in a cubic unit cell, with the common direction $[\bar{1}10]$.

Some useful conventions and results of the Miller notation are as follows:

- (i) Unknown Miller indices are denoted by symbols h , k and l . For example, for an unknown family of directions, we write $\langle hkl \rangle$.
- (ii) When the integers used in the Miller indices contain more than one digit, the indices must be separated by commas for clarity, e.g. (3, 10, 15).
- (iii) The crystal directions of a family are not necessarily parallel to one another. Similarly, not all members of a family of planes are parallel to one another.

- (iv) By changing the signs of all the indices of a crystal direction, we obtain the anti-parallel or opposite direction. By changing the signs of all the indices of a plane, we obtain a plane located at the same distance on the other side of the origin.
- (v) In cubic crystals, a crystal plane and a crystal direction normal to it have the same indices, e.g., $[11\bar{1}] \perp (11\bar{1})$.

As a review, some common directions and planes in cubic crystals are illustrated below:



The interplanar spacing between adjacent planes of Miller indices (hkl), d_{hkl} , is defined as the spacing between the first such plane and a parallel plane passing through the origin. For example, the interplanar spacing of (100) planes in cubic crystals is equal to the lattice parameter a . In x-ray studies, it becomes necessary to make a distinction between parallel sets of planes such as (100) and (200). For this purpose, the first (200) plane is taken to make an intercept of $1/2$ along the a -axis, $1/2$ being the reciprocal of the corresponding index. Hence, the interplanar spacing for (200) planes is only half of that for (100) planes. In cubic crystals, the following relationship gives the interplanar spacing as a function of the Miller indices and the lattice parameter:

$$d_{\{hkl\}} = \frac{a}{\sqrt{h^2 + k^2 + l^2}} \quad (3.2)$$

The interplanar spacing is the same for all members of a family of planes, as the indices are squared and summed in the denominator of Eq. (3.2).

Even though three noncoplanar vectors are sufficient to describe a plane or a direction in any crystal, a four-digit notation, $hkil$, called the Miller-Bravais indices, is used for hexagonal crystals. The use of such a notation enables crystallographically equivalent planes or directions in a hexagonal crystal to be denoted by the same set of indices. Three of the axes \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 are coplanar and lie on the basal plane of the hexagonal prism (see the unit cell in Table 3.1), with a 120° angle between them. The fourth axis is the \mathbf{c} axis perpendicular to the basal plane.

The determination of the Miller-Bravais indices is illustrated in Fig. 3.9 for the two examples given below. The indices of a prismatic plane (one of the vertical faces of the hexagonal prism) is of the type $(10\bar{1}0)$. It makes an intercept of 1 along a_1 , is parallel to a_2 axis, makes an intercept of -1 along a_3 axis and is again parallel to the c -axis. Note that, in this notation, $h + k = -i$. This equality always holds good for the Miller-Bravais indices. The indices of a direction that lies along one of the three axes (parallel to a side of the hexagon) is of the type $[2\bar{1}\bar{1}0]$. In the example given, the direction is parallel to the a_1 axis and is resolved into components along a_2 and a_3 . Each of these components being -1 , the value of the first index corresponding to the a_1 axis can be obtained as $-(-1 - 1) = 2$ from the condition that $h + k = -i$. This yields $[2\bar{1}\bar{1}0]$.

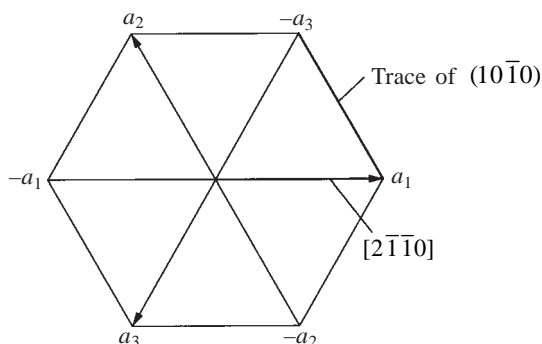


Fig. 3.9 Plan view of the hexagonal unit cell showing the plane $(10\bar{1}0)$ and the direction $[2\bar{1}\bar{1}0]$ in the Miller-Bravais notation.

STRUCTURE DETERMINATION BY X-RAY DIFFRACTION

It is well known that, for visible electromagnetic radiation to be diffracted, the spacing between lines in a two-dimensional grating must be of the same order as the wavelength range for light (3900–7800 Å). The same principle holds good for diffraction by the three-dimensional grating of the periodic array of atoms in crystals. The typical interatomic spacing in crystals is 2–3 Å. So, the wavelength of the radiation used for crystal diffraction should be in the same range. X-rays have wavelengths in this range and are, therefore, diffracted by crystals. This property is widely used for the study of crystal structures.

3.4 The Bragg Law of X-ray Diffraction

When electrons moving at high speeds are directed to a metal target, a small percentage of their kinetic energy is converted into x-rays. The x-rays emitted by the target consist of a continuous range of wavelengths, called *white radiation*, by analogy with white light consisting of a range of wavelengths. The minimum wavelength in the continuous spectrum is inversely proportional to the applied voltage which accelerates the electrons towards the target. If the applied voltage is sufficiently high, in addition to the white radiation, a *characteristic radiation* of a specific wavelength and high intensity is also emitted by the target. The radiation emitted by a molybdenum target at 35 kV includes both types of radiation as illustrated in Fig. 3.10. In spectroscopic notation, the characteristic

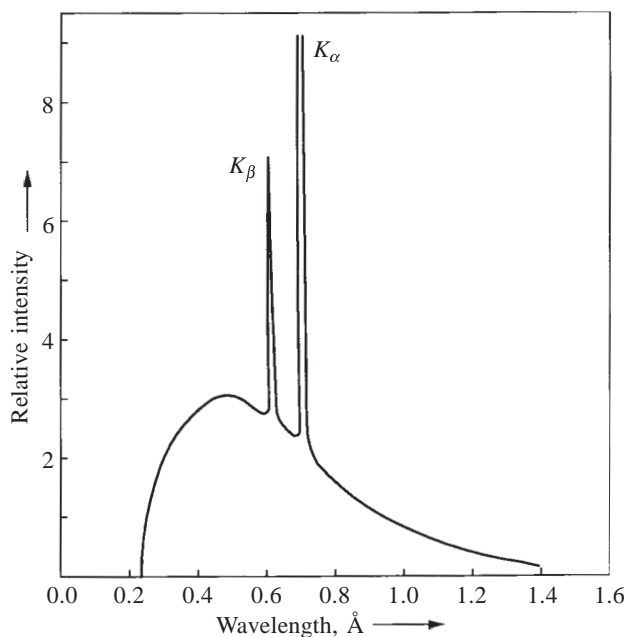


Fig. 3.10 The spectrum of x-rays emitted from a molybdenum target at 35 kV.

radiations are named K_α , K_β , L_α , etc. K_α radiation has a high intensity and is commonly used in diffraction studies. The wavelengths of this radiation for typical target metals are given in Table 3.2.

TABLE 3.2
Wavelengths of K_α Radiation for Typical Target Metals

Target metal	Mo	Cu	Co	Fe	Cr
K_α wavelength, Å	0.71	1.54	1.79	1.94	2.29
nm	0.071	0.154	0.179	0.194	0.229

A beam of x-rays directed at a crystal interacts with the electrons of the atoms in the crystal. The electrons oscillate under the impact and become a new source of electromagnetic radiation. The waves emitted by the electrons have the same frequency as the incident x-rays. The emission is in all directions. As there are millions of atoms in a crystal, the emission in a particular direction is the combined effect of the oscillations of electrons of all the atoms. The emissions will be in phase and reinforce one another only in certain specific directions, which depend on the direction of the incident x-rays, their wavelength as well as the spacing between atoms in the crystal. In other directions, there is destructive interference of the emissions from different sources. The easiest way to visualize the diffraction effects produced by the three-dimensional grating provided by the crystal is to consider the *Bragg law*.

In Fig. 3.11, a set of parallel planes in a crystal is shown. A beam of x-rays of wavelength λ is directed towards the crystal at an angle θ to the atomic

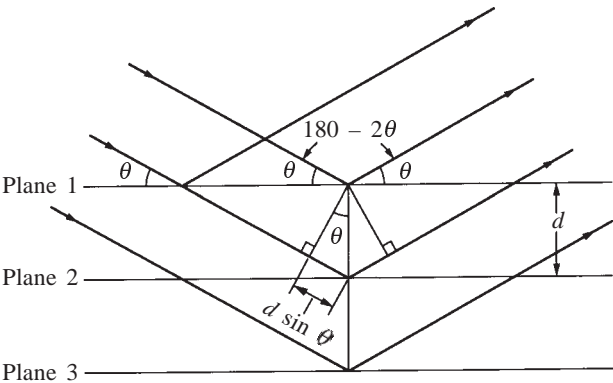


Fig. 3.11 **Illustration of the Bragg law.**

planes. In Bragg law, the interaction described above between x-rays and the electrons of the atoms is visualized as a process of reflection of x-rays by the atomic planes. This is an equivalent description of the diffraction effects produced by a three-dimensional grating. The atomic planes are considered to be

semi-transparent, that is, they allow a part of the x-rays to pass through and reflect the other part, the incident angle θ (called the Bragg angle) being equal to the reflected angle. Referring to Fig. 3.11, there is a path difference between rays reflected from plane 1 and the adjacent plane 2 in the crystal. The two reflected rays will reinforce each other, only when *this path difference is equal to an integral multiple of the wavelength*. If d is the interplanar spacing, the path difference is twice the distance $d \sin \theta$, as indicated in Fig. 3.11. The Bragg condition for reflection can therefore be written as

$$n\lambda = 2d \sin \theta \quad (3.3)$$

where n is an integer and λ is the wavelength of the x-rays used. A first order reflection is obtained, if $n = 1$; a second order reflection occurs if $n = 2$, and so on.

As $\sin \theta$ has a maximum value of 1, for a typical value of interplanar spacing of 2 Å, Eq. (3.3) gives the upper limit of λ for obtaining a first order reflection as 4 Å. There will be no reflection if λ is greater than 4 Å. λ can be reduced indefinitely, obtaining reflections from other sets of planes that have spacing less than 2 Å as well as an increasing number of higher order reflections. A very small wavelength of the order of 0.1 Å is not necessarily an advantage as it tends to produce other effects such as knocking off electrons from the atoms of the crystal and getting absorbed in the process. The wavelengths of the K_α radiation given in Table 3.2 for typical target metals lie in the right range.

The Bragg equation can be used for determining the lattice parameters of cubic crystals. Let us first consider the value that n should be assigned. A second order reflection from (100) planes should satisfy the following Bragg condition:

$$2\lambda = 2d_{100} \sin \theta$$

or

$$\lambda = d_{100} \sin \theta \quad (3.4)$$

Similarly, a first order reflection from (200) planes should satisfy the following condition:

$$\lambda = 2d_{200} \sin \theta \quad (3.5)$$

We have earlier noted that the interplanar spacing of (100) planes is twice that for (200) planes. So, Eqs. (3.4) and (3.5) are identical. For any incident beam of x-rays, the Bragg angle θ would be the same, as the two sets of planes in question are parallel. As Eqs. (3.4) and (3.5) are identical, the two reflections will superimpose on each other and cannot be distinguished. By a similar argument, it can be shown that the third order reflection from (100) planes will superimpose on the first order reflection from (300) planes. In view of such superimposition, there is no need to consider the variations in n separately; instead, we take n to be unity for all reflections from parallel sets of planes such as (100), (200), (300), (400), etc. In a crystal, it may turn out, for example, that there is no (200) plane with atoms on it. Then, what is designated as a (200) reflection actually refers to the second order reflection from (100) planes.

Example 3.7 A diffraction pattern of a cubic crystal of lattice parameter $a = 3.16 \text{ \AA}$ is obtained with a monochromatic x-ray beam of wavelength 1.54 \AA . The first four lines on this pattern were observed to have the following values:

Line	θ (in degrees)
1	20.3
2	29.2
3	36.7
4	43.6

Determine the interplanar spacing and the Miller indices of the reflecting planes.

Solution Using the Bragg equation, we can write the interplanar spacing $d = \lambda/(2 \sin \theta)$. n is assumed to be 1, as higher order reflections superpose on the lower order ones for parallel sets of planes. The d values can now be determined. Since $d = a/\sqrt{h^2 + k^2 + l^2}$ for cubic crystals, $h^2 + k^2 + l^2$ can also be determined. The results are tabulated below:

Line	$d_{hkl}, \text{ \AA}$	$(h^2 + k^2 + l^2) = a^2/d^2$	hkl
1	2.220	2	110
2	1.579	4	200
3	1.288	6	211
4	1.116	8	220

Starting from the lowest index plane, we notice that there are no reflections corresponding to $(h^2 + k^2 + l^2) = 1, 3$ and 5 . There is no plane which has $(h^2 + k^2 + l^2) = 7$. Corresponding to $(h^2 + k^2 + l^2) = 2, 4, 6$ and 8 , there are reflections from $\{110\}$, $\{200\}$, $\{211\}$ and $\{220\}$ planes, respectively.

In diffraction studies, in order to increase the probability that crystals with the right orientation for Bragg reflection are available, one of the following procedures is adopted:

- A monochromatic x-ray beam of a specific wavelength is combined with numerous possible θ values so that reflection occurs at the right combination that satisfies the Bragg law. This is done by placing thousands of crystals of random orientation in the path of the beam. The crystals are usually in *powder form*.
- A single crystal is held stationary in the path of the beam so that θ is kept constant. A white radiation is then directed at the crystal so that numerous values of the wavelength are available, and again the right combination will lead to the diffraction condition. This method is called the *Laue technique*.
- A single crystal is held in the beam of a monochromatic radiation and is rotated such that at some position of the crystal, the diffraction condition is satisfied. This method is known as *the rotating crystal method*. Even though this is not the most widely used method, it provides greater certainty in identification, as well as more accurate measurement of the intensities of the reflected beam.

3.5 The Powder Method

The powder method is a widely used experimental technique for the routine determination of crystal structures. It is highly suitable for identification and for determination of the structures of crystals of high symmetry. Here, a monochromatic x-ray beam, usually of K_α radiation, is incident on thousands of randomly oriented crystals in powder form. The powder camera, called the *Debye-Scherrer camera*, consists of a cylindrical cassette, with a strip of photographic film positioned around the circular periphery of the cassette. The powder specimen is placed at the centre of the cassette in a capillary tube or pasted on a thin wire. The tube, the wire and the paste material must be of some nondiffracting substance such as glass or glue. The x-ray beam enters through a small hole, passes through the powder specimen and the unused part of the beam leaves through a hole at the opposite end. The geometry of the powder method is illustrated in Fig. 3.12.

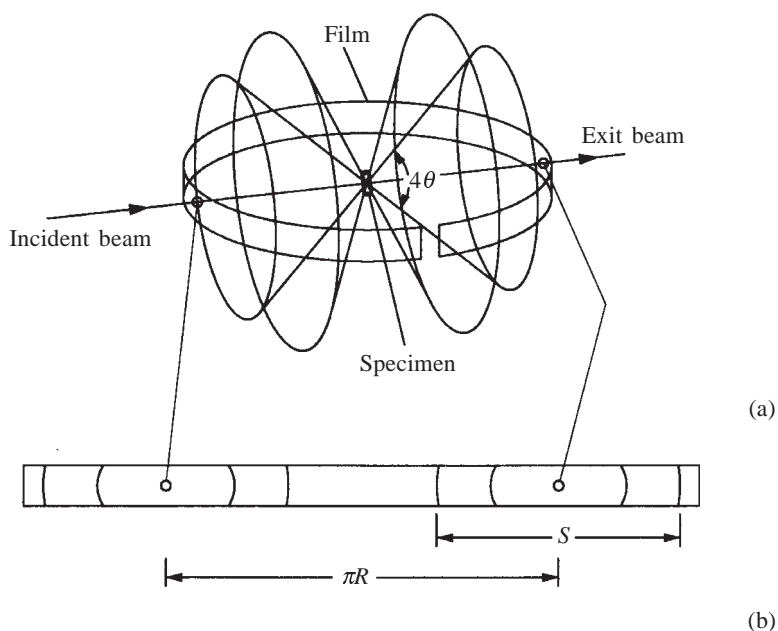


Fig. 3.12 The geometry of the powder method.

Consider a set of parallel crystal planes making an angle θ with the incident direction. When this angle satisfies the Bragg equation, there is reflection. By virtue of the large number of randomly oriented crystals in the powder, there are a number of possible orientations of this set of planes in space for the same angle θ with the incident direction. So the reflected radiation is not just a pencil beam like the incident one; instead, it lies on the surface of a cone whose apex is at the point of contact of the incident radiation with the specimen. Also, the interplanar spacing d being the same for all members of a family of crystal

planes, they all reflect at the same Bragg angle θ , all reflections from a family lying on the same cone. After taking $n = 1$ in the Bragg equation, there are still a number of combinations of d and θ that would satisfy the Bragg law. For each combination of d and θ , one cone of reflection must result and, therefore, many cones of reflection are emitted by the powder specimen. If the reflected cones were recorded on a flat film placed normal to the exit beam, they will be in the form of concentric circles. In the powder camera, however, only a part of each reflected cone is recorded by the film strip positioned at the periphery of the cylindrical cassette. *The recorded lines from any cone are a pair of arcs that form part of the circle of intersection.* When the film strip is taken out of the cassette and spread out, it looks like Fig. 3.12b.

Note that the angle between a reflected line lying on the surface of the cone and the exit beam is 2θ . Therefore, the angle included at the apex of the cone is twice this value, 4θ , Fig. 3.12a. When the Bragg angle is 45° , the cone opens out into a circle and reflection at this angle will make a straight line intersection with the film strip at the midpoint between the incident and the exit points in Fig. 3.12b. When the Bragg angle is greater than 45° , back reflection is obtained, that is, the reflected cones are directed towards the incident beam. Bragg angles up to the maximum value of 90° can be recorded by the film of the powder camera, which is not possible on a flat film placed in front of the exit beam.

The exposure in a powder camera must be sufficiently long to give reflected lines of good intensity. The exposure time is usually a few hours. After the film is exposed and developed, it is indexed to determine the crystal structure. It is easily seen that the first arc on either side of the exit point corresponds to the smallest angle of reflection. The pairs of arcs beyond this pair have larger Bragg angles and are from planes of smaller spacings, recall that $d = \lambda/(2 \sin \theta)$. The distance between any two corresponding arcs on the spread out film is termed S , Fig. 3.12b. S is related to the radius of the powder camera R :

$$S = 4R\theta \quad (3.6)$$

where θ is the Bragg angle expressed in radians. For easy conversion of the distance S measured in mm to Bragg angle in degrees, the camera radius is often chosen to be 57.3 mm, as $1 \text{ rad} = 57.3^\circ$.

In the powder method, the intensity of the reflected beam can also be recorded in a diffractometer, which uses a counter in place of the film to measure intensities. The counter moves along the periphery of the cylinder and records the reflected intensities against 2θ . Peaks in the diffractometer recording (Fig. 3.13) correspond to positions where the Bragg condition is satisfied.

3.6 Structure Determination

The determination of a complex crystal structure is often time consuming, requiring a lot of patience and ingenuity. A step-by-step procedure is followed in such cases, first determining the macroscopic symmetry of the crystal, then the space lattice and its dimensions, and finally the atomic arrangement within the unit cell. Measurement of the density of the crystal and the chemical composition also assist the process of structure determination.

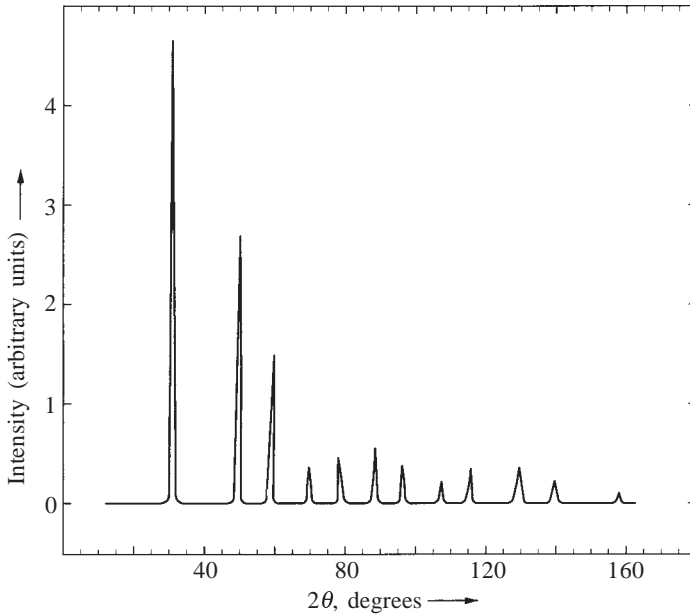


Fig. 3.13 The tracing from a diffractometer.

In simple crystals of high symmetry, the space lattice and its dimensions can be determined relatively easily. If the crystal is monoatomic, the space lattice together with the lattice parameters is a complete description of the crystal structure. If, on the other hand, the basis is two or more atoms per lattice point, the number and distribution of atoms within the unit cell can be determined only from quantitative measurements of the reflected intensities. For such measurements, the recording from a diffractometer is more useful than the pattern obtained from a powder camera.

The procedure for determining the structure of monoatomic cubic crystals is outlined below. Combining Eq. (3.2) for the interplanar spacing d with the Bragg equation, we obtain

$$\sin^2 \theta = \frac{\lambda^2}{4a^2} (h^2 + k^2 + l^2) \quad (3.7)$$

n is assumed to be 1 for reasons already outlined. θ -values can be determined from a powder pattern using Eq. (3.6). Since monochromatic radiation is used in the powder technique, the value of λ is known. Then, the unknowns in Eq. (3.7) are the Miller indices of the reflecting planes that correspond to the measured angles of reflection. For a given cubic lattice, it is possible to list all combinations of h , k and l and arrange $(h^2 + k^2 + l^2)$ in increasing order which will also be the increasing order of θ values, as seen from Eq. (3.7). The $\sin^2 \theta$ values will be in the same ratio as $(h^2 + k^2 + l^2)$, if the assumed and actual lattices coincide.

The distinction between lattices of the cubic system is possible by using the fact that not all combinations of $(h^2 + k^2 + l^2)$ lead to reflection for a given lattice. Consider, for example, the first order reflection from the (100) planes of a BCC crystal. These planes are the faces of the unit cube and contain the ‘corner’ atoms of the cube, see Fig. 3.14. The path difference between reflected beams from two adjacent (100) planes is one full wavelength and, therefore, the reflected beams are in phase. The midway parallel plane between these two (100) planes contains the body centred atoms. It is easily seen that the reflection from this midplane will be out of phase by exactly half of a wavelength with the (100) reflections, see Fig. 3.14. As the effective number of body centred atoms is equal to the effective number of corner atoms in a BCC crystal, the intensity of the reflected beams

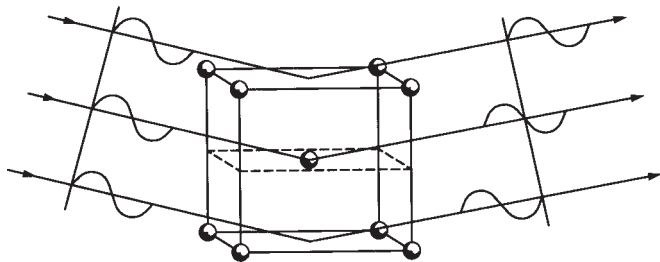


Fig. 3.14 The (100) reflection is absent for a BCC crystal, as reflection from the corner atoms is exactly cancelled out by that from body centred atoms.

from atoms at these two locations will be exactly equal. The phase difference $\lambda/2$ then results in a net zero reflected intensity. (Note that the terms, body centred atoms and body corner atoms, are defined only in a relative sense and are interchangeable.) There is thus no first order reflection from (100) planes in a BCC crystal. A second order reflection from (100) planes is possible, but this will superimpose on the first order reflection from (200) planes.

By following a similar reasoning, it is possible to derive extinction rules for different cubic crystals, as given in Table 3.3.

TABLE 3.3
Extinction Rules for Cubic Crystals

Crystal	Reflections are allowed
SC	for all values of $(h^2 + k^2 + l^2)$
BCC	for even values of $(h + k + l)$
FCC	when h, k and l are all odd or all even
DC	when h, k and l are all odd, or when all are even, $(h + k + l)$ should be divisible by four

In the above, zero is taken as an even number. The diamond cubic (DC) crystal is based on the FCC space lattice, with a basis of two atoms per lattice point. The DC structure is discussed in Chap. 5.

From the extinction rules, we can derive the ratio of $(h^2 + k^2 + l^2)$ values for allowed reflections from different crystals:

SC	1 : 2 : 3 : 4 : 5 : 6 : 8 ...
BCC	1 : 2 : 3 : 4 : 5 : 6 : 7 ...
FCC	3 : 4 : 8 : 11 : 12 ...
DC	3 : 8 : 11 : 16 ...

A simple comparison of the observed ratios of $\sin^2 \theta$ values with the above is then sufficient to identify the crystal structure.

Example 3.8 From a powder camera of diameter 114.6 mm, using an x-ray beam of wavelength 1.54 Å, the following S values in mm are obtained for a material:

86, 100, 148, 180, 188, 232, and 272

Determine the structure and the lattice parameter of the material.

Solution As the camera radius $R = 57.3$ mm, the Bragg angles in degrees at which reflections are observed are equal to $S/4$:

21.5, 25, 37, 45, 47, 58 and 68.

The $\sin^2 \theta$ values are in the ratio

0.1346 : 0.1788 : 0.3620 : 0.5003 : 0.5352 : 0.7195 : 0.8596

Within experimental error, these values can be expressed in the ratio of integral numbers

3 : 4 : 8 : 11 : 12 : 16 : 19

From the extinction rules, we note that the structure is FCC. The lattice parameter calculated from the highest Bragg angle is 3.62 Å. By referring to table on back inside cover, we note that the material is likely to be copper.

SUMMARY

1. A space lattice is an infinite array of points, all with identical surroundings.
2. A crystal structure is obtained by combining a space lattice with a basis. The basis must give the number of atoms per lattice point, their types, mutual orientations and distances of separation.

3. Space lattices are limited to fourteen, but known crystal structures run into thousands.
4. Crystal directions and crystal planes are denoted by the Miller indices. A family of crystal directions or planes includes all possible combinations of the indices, both positive and negative.
5. The Bragg law gives the condition for diffraction by a crystal.
6. The powder method is the commonly used x-ray diffraction technique.
7. From the positions of lines in a powder pattern and from the extinction rules for different cubic crystals, the space lattice and its dimensions can be determined.

PROBLEMS

- 3.1** Draw two adjacent FCC unit cells and join the top and bottom face centres of each cell. Also join these four face centres to the nearest common corners.
- (i) What is the unit cell outlined by this procedure?
 - (ii) Is it in the Bravais list?
 - (iii) If so, the FCC cell with a higher symmetry must be a special case of the outlined cell. Explain.
- 3.2** There is no end-centred tetragonal lattice in the Bravais list, but there is an end-centred orthorhombic lattice. Explain why this is so.
- 3.3** Draw a:
- (i) (111) plane in a triclinic lattice;
 - (ii) $(1\bar{2}0)$ plane in a tetragonal lattice;
 - (iii) $[11\bar{1}]$ direction in an orthorhombic lattice; and
 - (iv) $[\bar{2}11]$ direction in a cubic lattice.
- 3.4** Draw a (111) plane and a (222) plane in the unit cell of a cubic lattice with lattice parameter a . Determine their distances from a parallel plane through the origin.
- Answer:* $a/\sqrt{3}$ and $a/(2\sqrt{3})$.
- 3.5** Find the Miller indices of a plane that makes an intercept of 1 on the a -axis and 2 on the b -axis and is parallel to the c -axis.
- Answer:* (210).
- 3.6** Determine the Miller indices of a plane that makes an intercept of 2 Å, 3 Å and 4 Å on the coordinate axes of an orthorhombic crystal with $a : b : c = 4 : 3 : 2$.
- Answer:* (421).

- 3.7** Find the Miller indices of a plane that makes intercepts on a , b and c axes equal to 3 Å, 4 Å and 3 Å in a tetragonal crystal with the c/a ratio of 1.5.
Answer: (436).

- 3.8** Draw a $(11\bar{1})$ plane in the unit cell of a cubic crystal. Show all the $\langle 110 \rangle$ directions that lie on this plane, giving the Miller indices of each one of them.

Answer: $[\bar{1}10]$, $[0\bar{1}\bar{1}]$, $[101]$ and the three opposite directions $[1\bar{1}0]$, $[011]$, $[\bar{1}0\bar{1}]$.

- 3.9.** Sketch a (110) plane in the unit cell of a cubic crystal. Show all the $\langle 111 \rangle$ directions that lie on this plane, giving the Miller indices of each one of them.

Answer: $[1\bar{1}1]$, $[\bar{1}11]$ and $[\bar{1}\bar{1}\bar{1}]$, $[1\bar{1}\bar{1}]$.

- 3.10** What are the Miller indices of the line of intersection of a $(\bar{1}\bar{1}1)$ and a $(\bar{1}\bar{1}\bar{1})$ plane in a cubic crystal? Determine the answer both geometrically and analytically.

Answer: $[1\bar{1}0]$ or $[\bar{1}10]$.

- 3.11** Calculate the atomic density (number of atoms per unit area) in (111) , (110) and (100) planes of copper (FCC) with the lattice parameter of 3.61 Å. Can you pack atoms more closely than in (111) plane?

Answer: 1.77×10^{19} , 1.08×10^{19} and 1.53×10^{19} atoms m^{-2} .

- 3.12** Calculate the number of atoms per unit area of (111) , (110) and (100) planes of a BCC crystal. The answer may be derived as a function of the lattice parameter a .

Answer: $1/(a^2\sqrt{3})$, $\sqrt{2}/a^2$ and $1/a^2$.

- 3.13** Express the edge, face diagonal and body diagonal of the unit cell in terms of the atomic radius r for SC, BCC and FCC crystals.

- 3.14** The distance between (111) planes in a face centred cubic crystal is 2 Å. Determine the lattice parameter and the atomic diameter.

Answer: $2\sqrt{3}$ Å, $\sqrt{6}$ Å.

- 3.15** A BCC crystal is used to measure the wavelength of some x-rays. The Bragg angle for reflection from (110) planes is 20.2° . What is the wavelength? The lattice parameter of the crystal is 3.15 Å.

Answer: 1.54 Å.

- 3.16** X-rays with a wavelength of 1.54 Å are used to calculate the spacing of (200) planes in aluminium. The Bragg angle for this reflection is 22.4° . What is the size of the unit cell of the aluminium crystal?

Check your answer with table on back inside cover.

- 3.17** Determine the structure of the crystal in Example 3.7.

Answer: BCC.

- 3.18** Using a diffractometer and a radiation of wavelength 1.54 \AA , only one reflection from an FCC material is observed when 2θ is 121° . What are the indices of this reflection? What is the interplanar spacing? Show that the next higher index reflection cannot occur.

Answer: $\{111\}$, 0.885 \AA .

- 3.19** The first three lines from the powder pattern of a cubic crystal have the following S values: 24.95, 40.9 and 48.05 mm. The camera radius is 57.3 mm. Molybdenum K_α radiation of wavelength 0.71 \AA is used. Determine the structure and the lattice parameter of the material.

Answer: DC, 5.66 \AA .

- 3.20** The Bragg angle corresponding to a reflection for which $(h^2 + k^2 + l^2) = 8$ is found to be 14.35° . Determine the lattice parameter of the crystal. X-rays of wavelength 0.71 \AA are used. If there are two other reflections with smaller Bragg angles, what is the crystal structure?

Answer: 4.05 \AA , FCC.

- 3.21** Aluminium (FCC) has a lattice parameter of 4.05 \AA . When a monochromatic radiation of 1.79 \AA is used in a powder camera, what would be the first four S values? The camera diameter is 114.6 mm.

Answer: 90.1, 105.0, 154.7, and 188.4 mm.

- 3.22** The first reflection using copper K_α radiation from a sample of copper powder (FCC) has an S value of 86.7 mm. Compute the camera radius.

Answer: 57.3 mm.

- 3.23** Copper and nickel (both FCC) dissolve in each other in all proportions at room temperature in the solid state. From a sample of Cu–Ni alloy of unknown composition, a powder pattern is obtained using monochromatic radiation of wavelength 1.54 \AA . The spread-out film showed that, for the first reflection, $S = 88.4$ mm. The distance between inlet and outlet points of the x-ray beam was 180 mm. Estimate the percentage of copper in the alloy, assuming a linear variation of lattice parameter with composition.

Answer: 28 atomic %.

- 3.24** The diffractometer tracing shown in Fig. 3.13 is of silicon (DC). Index the reflecting planes corresponding to each peak.

- 3.25** Name the space lattices, which have four lattice parameters.

- 3.26** In an orthorhombic unit cell, $a : b : c = 1 : 2 : 3$. The magnitude of a is 2 \AA . What are the intercepts in \AA of a plane of Miller indices (230)?

Answer: 1 \AA , 1.33 \AA , ∞ .

- 3.27** In cubic crystals, $\{110\}$ and $\{112\}$ planes have a common $\langle 111 \rangle$ type of direction. Within a cubic unit cell, sketch one plane of each type such that a member of the above family of directions is common to both the planes. Indicate in your sketch the Miller indices of the two planes and the line of intersection.

- 3.28** Explain why there is no end-centred cubic space lattice.

- 3.29** Explain with a neat sketch why a (100) reflection is not possible from an FCC crystal.
- 3.30** A Debye-Scherrer pattern is obtained from an FCC crystal with $a = 3.61 \text{ \AA}$. X-rays of wavelength equal to 1.54 \AA are used. Determine the Miller indices of the planes with the lowest and the highest Bragg angles in this pattern.
Answer: {111} and {420}.
- 3.31** Only two reflections are observed on the powder pattern of a cubic crystal at Bragg angles of 34.06° and 65.51° . Chromium K_α radiation is used. Determine the crystal structure (if possible) and the lattice parameter. Show that the third and higher reflections cannot occur.
Answer: DC, 3.56 \AA .
- 3.32** The entry point and the exit point of x-rays on a powder pattern taken from a cubic material could not be distinguished. Assuming one of the points to be the exit point, the following S values were obtained: 311.95, 319.10 and 335.05 mm. The camera radius is 57.3 mm. Molybdenum K_α radiation was used. Determine the structure and the lattice parameter of the material.
Answer: DC, 5.66 \AA .
- 3.33** The first two reflections as seen in the powder photograph of a cubic material have S values of 90.1 and 154.7 mm, respectively. As the film between these two reflections was torn, it was not clear whether an additional line was present between the two observed lines. Determine, if possible, the crystal structure and the lattice parameter. The third and the fourth reflections as seen in the torn film have S values of 188.4 and 199.8 mm. The camera radius is 57.3 mm. Cobalt K_α radiation was used.
Answer: FCC, 4.05 \AA .
- 3.34** In the damaged powder film of a cubic material, the locations of only the first two pairs of lines could be accurately measured. They had S values of 24.95 and 40.9 mm. The camera radius is 57.3 mm. Molybdenum K_α radiation was used. If possible, determine the crystal structure and the lattice parameter of the material. The third pair of lines in the damaged film was estimated to have an S value of about 49 mm. Determine whether this confirms your conclusion or not.
Answer: DC, 5.66 \AA .

MULTIPLE CHOICE QUESTIONS

- The space lattices with two lattice parameters belong to the crystal systems
 A. tetragonal B. rhombohedral C. hexagonal D. triclinic

2. The number of lattice points in the rhombohedral unit cell is
A. 8 B. 4 C. 2 D. 1
3. There is no end-centred orthorhombic lattice
A. as it violates the definition of a space lattice
B. as it can be represented by a simple orthorhombic unit cell
C. as it can be represented by a simple monoclinic unit cell
D. the question is wrong
4. The effective number of lattice points in the unit cell of end-centred monoclinic lattice is
A. 1 B. 2 C. 4
D. can't be found unless the angle β is given
5. There is no end-centred cubic space lattice, because
A. it can be represented by the simple cubic lattice
B. it can be represented by the simple orthorhombic lattice
C. it violates the cubic symmetry
D. none of these
6. The unit cell with three lattice parameters is
A. tetragonal B. orthorhombic C. monoclinic D. triclinic
7. A unit cell has $a = 5 \text{ \AA}$, $b = 8 \text{ \AA}$, $c = 3 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 65^\circ$ and $\gamma = 54^\circ$.
The space lattice for this unit cell is
A. orthorhombic B. monoclinic C. rhombohedral D. triclinic
8. The number of Bravais space lattices with two lattice points are
A. 2 B. 3 C. 4 D. 5
9. The tetragon has
A. 4 faces B. 12 edges C. 6 corners D. 8 edges
10. The minimum number of ions in the unit cell of an ionic crystal with FCC space lattice is
A. 4 B. 8 C. 12 D. 16
11. The atomic diameter of an FCC crystal (lattice parameters is a) is
A. $a\sqrt{2}/2$ B. $a\sqrt{2}/4$ C. $a\sqrt{3}/4$ D. $a/2$
12. If the radius of an atom in a simple cubic crystal is r , the body diagonal of the unit cell is
A. $r\sqrt{3}$ B. $2r\sqrt{3}$ C. $4r/\sqrt{3}$ D. $3r/4$
13. The number of unit cells in 1 m^3 of FCC nickel ($r_{\text{Ni}} = 1.243 \text{ \AA}$) is
A. 2.3×10^{28} B. 4.2×10^{28} C. 6.5×10^{28} D. 18.4×10^{28}
14. The acute angle between $[101]$ and $[\bar{1}01]$ directions in a tetragonal crystal with $c/a = 1.5$ is
A. 90° B. 67.38° C. 56.30° D. 33.69°

15. The angle between $[111]$ and $[11\bar{2}]$ directions in a cubic crystal is
A. 0° B. 45° C. 90° D. 180°
16. The number of members in the family $\langle 123 \rangle$ in a cubic crystal are
A. 8 B. 12 C. 24 D. 48
17. The $(1\bar{1}1)$ plane is parallel to
A. (111) B. $(\bar{1}\bar{1}\bar{1})$ C. $(\bar{1}\bar{1}1)$ D. $(11\bar{1})$
18. The Miller indices of a plane in an orthorhombic unit cell ($a : b : c = 1 : 2 : 3$) making intercepts of 3, 2, 1 Å on a , b , c , axes are
A. (931) B. (139) C. (321) D. (123)
19. The four non-parallel body diagonals of a cube have Miller indices
A. $[111]$, $[\bar{1}\bar{1}1]$, $[1\bar{1}\bar{1}]$, $[11\bar{1}]$ B. $[111]$, $[\bar{1}\bar{1}\bar{1}]$, $[1\bar{1}\bar{1}]$, $[\bar{1}\bar{1}1]$
C. $[\bar{1}\bar{1}\bar{1}]$, $[1\bar{1}\bar{1}]$, $[\bar{1}\bar{1}1]$, $[\bar{1}\bar{1}1]$ D. $[\bar{1}\bar{1}1]$, $[\bar{1}\bar{1}\bar{1}]$, $[1\bar{1}\bar{1}]$, $[\bar{1}\bar{1}1]$
20. The Miller indices of the line of intersection of a $(1\bar{1}1)$ and a $(1\bar{1}0)$ plane are
A. $[1\bar{1}0]$ B. $[110]$ C. $[\bar{1}\bar{1}0]$ D. $[1\bar{1}1]$
21. If the interplanar spacing obtained from the second reflection of a DC crystal is 1.81 Å, the lattice parameter is
A. 0.905 Å B. 2.56 Å C. 3.62 Å D. 5.12 Å
22. The first S value on the powder pattern of an FCC crystal ($a = 4.05$ Å) taken with a camera of radius 57.3 mm using cobalt K_α radiation of 1.79 Å is
A. 45 mm B. 51 mm C. 72.8 mm D. 90.0 mm.
23. The second order reflection from (200) planes coincides with the first order reflection from
A. (100) B. (200) C. (400) D. none of these
24. The interplanar spacing of the first reflecting plane (lowest θ) in an FCC crystal
A. $a\sqrt{3}$ B. $a/\sqrt{3}$ C. $a/\sqrt{2}$ D. a
25. The first reflection from polonium powder (SC) occurs at Bragg angle of 28.7° . The fourth reflection is at Bragg angle
A. 57.4° B. 114.8° C. 73.8° D. none of these
26. If the first reflection from a BCC crystal has a Bragg angle θ of 22.2° , the second reflection will have a θ of
A. 32.3° B. 25.9° C. 38.1° D. 44.4°
27. No reflection will be observed from a DC crystal ($a = 3$ Å) if the wavelength of x-rays is greater than
A. 3.0 Å B. 3.25 Å C. 3.37 Å D. 3.46 Å

28. If the first reflection from a FCC crystal has a Bragg angle θ of 21.5° , the second reflection will have an angle θ of
 A. 18.5° B. 25° C. 31.2° D. 36.8°
29. The first three reflecting planes of silicon (DC) are
 A. 111, 200, 220 B. 110, 200, 211
 C. 111, 220, 311 D. 100, 110, 111
30. Using X-rays of 2.29 \AA , the first two reflections have Bragg angles of 34.06° and 65.51° . The Bragg angle for the third reflection is
 A. 68.1° B. 75.9° C. 89.9° D. reflection absent
31. The Miller indices of the fifth reflection in an FCC crystal is
 A. 331 B. 222 C. 311 D. 400

Answers

- | | | | | |
|------------|-------|-------|----------|----------|
| 1. A, B, C | 2. D | 3. D | 4. B | 5. C |
| 6. B | 7. D | 8. D | 9. B | 10. B |
| 11. A | 12. B | 13. A | 14. B | 15. C |
| 16. D | 17. B | 18. B | 19. A, C | 20. B, C |
| 21. D | 22. D | 23. C | 24. B | 25. C |
| 26. A | 27. D | 28. B | 29. C | 30. D |
| 31. B | | | | |

Source for Experimental Data

P. Villars and L.D. Calvert, *Pearson's Handbook of Crystallographic Data for Intermetallic Phases*, Vols. 1–3, American Society for Metals, Metals Park, Ohio (1985).

Suggestions for Further Reading

- B.D. Cullity, *Elements of X-ray Diffraction*, Addison Wesley, Reading, Mass. (1978).
- C. Hammond, *The Basics of Crystallography and Diffraction*, Oxford University Press, Oxford (2001).

Atomic Structure and
Chemical Bonding

The rigidity of a solid arises from the fact that the atoms in the solid are held together by interatomic bonds. The spatial arrangement of atoms in a solid is strongly influenced by the nature of these bonds, which in turn is influenced by the electronic structure of the atoms. The topic of chemical bonding is usually covered in the first course on general chemistry. Here, we shall briefly review the structure of the atom in relation to the periodic table and then the chemical bonding, as relevant to our discussion in the later chapters. The general variation of properties with bonding character is discussed at the end of the chapter.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Frequency of radiation ν	per second or hertz	s^{-1} or Hz	—
Momentum	kilogram metre per second	kg m s^{-1}	—
Electron energy level	joule	J	eV (electron volt), erg
Ionization potential Electron affinity Bond energy	kilojoule per mole	kJ mol^{-1}	eV/atom, kcal/mole
Bond length	nanometre	nm	Å
Dipole moment	coulomb metre	C m	debye

Constants

Planck's constant	$h = 6.626 \times 10^{-34} \text{ J s}$
Electronic charge	$e = 1.602 \times 10^{-19} \text{ C}$
Electron rest mass	$m_0 = 9.109 \times 10^{-31} \text{ kg}$
Velocity of light	$c = 2.998 \times 10^8 \text{ m s}^{-1}$

STRUCTURE OF THE ATOM

As the reader is aware, the atom consists of a nucleus and surrounding electrons. The nucleus is composed of protons and neutrons. As the mass of the electron is negligible compared to that of protons and neutrons, the mass of the atom depends mostly on the number of protons and neutrons in the nucleus. The neutrons carry no charge, the protons are positively charged and the electrons carry a negative charge. The charge on a proton or an electron is equal to $1.602 \times 10^{-19} \text{ C}$.

4.1 The Quantum States

The electrons surrounding the nucleus in an atom occupy different *orbitals*. The Heisenberg uncertainty principle states that the momentum and the position of an electron cannot be specified precisely. The product of the uncertainty in the momentum Δp and the uncertainty in the position Δx cannot be less than a certain value:

$$\Delta p \Delta x \geq \frac{h}{2\pi} \quad (4.1)$$

where h is Planck's constant. An outcome of this principle is that we cannot visualize an electron orbital to be a discrete path around the nucleus. Instead, we think of the orbital as an *electron probability density cloud* surrounding the nucleus.

Each electron orbital is called a *quantum state* with a set of quantum numbers n , l , m_l and m_s assigned to it. The characteristics of these quantum numbers are summarized as follows:

- (i) n is called the *principal quantum number* as it defines a principal electron orbital. n takes integer values 1, 2, 3, 4, 5, The larger the value of n , the farther removed is the orbital from the nucleus.
- (ii) l is called the *orbital angular momentum quantum number*. It is related to the shape of the electron orbital. l can take integer values from $l = 0$ to $l = n - 1$. If $n = 1$, $l = 0$. If $n = 2$, $l = 0$ or 1, and so on. The value of l defines suborbitals, which are denoted by certain letters:

$$l = 0, 1, 2, 3, 4$$

notation: s, p, d, f, g

- (iii) m_l is the *magnetic quantum number* and defines the spatial orientation of the electron probability density cloud. m_l can take values: $-l, -l + 1, \dots, -1, 0, +1, \dots, l - 1$, and l , so that there are $(2l + 1)$ values of m_l .
- (iv) m_s is called the *electron spin quantum number*. It defines the spin of the electron. There are two possible spins: spin up \uparrow and spin down \downarrow , corresponding to $m_s = +1/2$ and $-1/2$.

The *Pauli exclusion principle* states that no two electrons can have the same set of quantum numbers. In other words, only one electron can occupy a given quantum state. When there are a number of electrons surrounding the nucleus of an atom, they have to occupy different quantum states. In the lowest energy state of the atom known as the *ground state*, the electrons occupy the lowest energy levels, without at the same time violating the Pauli exclusion principle. The order of increasing energy of the orbitals is as follows:

1s
2s 2p
3s 3p
4s 3d 4p
5s 4d 5p
6s 4f 5d 6p
7s

In the above, the digit gives the value of the principal quantum number n . The following letter denotes the suborbital corresponding to different values of l .

An electron can be excited from a lower energy state to a higher energy vacant state by the supply of energy from an external source. When there is such an electronic transition from one level to another, electromagnetic radiation of a specific frequency and wavelength is absorbed or emitted. Energy is absorbed, if the electron is excited to a higher energy level. It is emitted if the electron is falling to a lower energy level from a higher level. The energy difference ΔE between the two states is related through the Einstein relationship to the frequency ν of the radiation emitted or absorbed:

$$\Delta E = h\nu \quad (4.2)$$

where h is Planck's constant. We can relate the frequency of the radiation to its wavelength λ through the velocity of light c :

$$c = \nu\lambda \quad (4.3)$$

4.2 The Periodic Table

The periodic table of elements is given in the front inside cover of the book. The following discussion is best understood by frequent reference to that table. Let us start from the first element, hydrogen, the lightest known atom. The atomic

number Z of hydrogen is 1, equal to the number of protons in the nucleus or the number of electrons surrounding the nucleus in the neutral atom. When the hydrogen atom is in the ground state, the electron occupies the $1s$ orbital immediately surrounding the nucleus. This electron configuration is denoted by $1s^1$.

The electron probability cloud of an s -orbital is spherically symmetric, as shown in Fig. 4.1a. The electron probability density ρ decreases with increasing distance from the nucleus. The radial density, $4\pi r^2\rho$ measures the probability of finding the electron in a thin spherical shell of radius r . The radial density is zero at the nucleus. It increases to a maximum value at a radial distance r_0 from the nucleus, where the electron is most likely to be found, see Fig. 4.1b. Thereafter, the radial density falls with further increase in the radial distance.

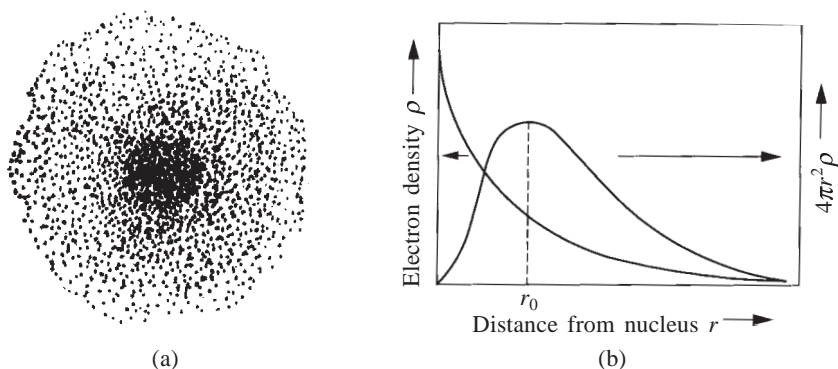


Fig. 4.1 (a) The electron probability cloud of hydrogen. (b) The electron density as a function of distance from nucleus.

The energy of the $1s$ electron is equal to $-2.18 \times 10^{-18} \text{ J}$ (-13.6 eV). The reference state of zero energy is that in which the electron is at an infinite distance from the nucleus. The negative value of the energy indicates that the system of the proton and the electron has lost energy to the surroundings, as the electron was *attracted* to the $1s$ orbital from infinity. The $2s$ orbital of the hydrogen atom (which is normally vacant) has a higher energy equal to $-5.44 \times 10^{-19} \text{ J}$ (-3.4 eV).

Example 4.1 Calculate the frequency and the wavelength of the radiation emitted, when an electron falls from the $2s$ level to the $1s$ level of the hydrogen atom.

Solution The energy difference between the two levels ΔE is

$$2.18 \times 10^{-18} - 5.44 \times 10^{-19} = 1.64 \times 10^{-18} \text{ J}$$

From the Einstein relation (Eq. (4.2)),

$$\begin{aligned}\text{Frequency of radiation emitted, } \nu &= 1.64 \times 10^{-18} / (6.626 \times 10^{-34}) \\ &= 2.48 \times 10^{15} \text{ Hz}\end{aligned}$$

From Eq. (4.3),

$$\begin{aligned}\text{Wavelength of radiation emitted, } \lambda &= 2.998 \times 10^8 / (2.48 \times 10^{15}) \\ &= 1.21 \times 10^{-7} \text{ m} \\ &= 1210 \text{ \AA}\end{aligned}$$

The element next to hydrogen is the inert gas helium with an atomic number $Z = 2$. It has two protons and two neutrons in its nucleus. The charge of the protons is balanced by two electrons in the $1s$ orbital. When $n = 1$,

$$l = 0, \quad m_l = 0, \quad m_s = \pm 1/2$$

Thus, only two quantum states are possible for $n = 1$. The two electrons of the $1s$ orbital are of opposite spins and have spin quantum numbers $m_s = +1/2$ and $-1/2$. When the helium atom is in the ground state, the $2s$ orbital is vacant as in the case of hydrogen.

When $n = 2$, we can have

$$\left. \begin{array}{lll} l = 0, & m_l = 0, & m_s = \pm 1/2 \\ l = 1, & m_l = -1, & m_s = \pm 1/2 \\ & m_l = 0, & m_s = \pm 1/2 \\ & m_l = +1, & m_s = \pm 1/2 \end{array} \right\} \begin{array}{l} 2s \\ 2p \end{array}$$

Therefore, the maximum number of electrons in the second principal orbital is eight. Correspondingly, there are eight elements in the second row of the periodic table. The first element is the alkali metal lithium with $Z = 3$. It has two electrons of opposite spin in the $1s$ orbital and a third electron in the $2s$ orbital. This filling order is in accord with the Pauli exclusion principle and the minimum energy criterion. This electronic configuration is denoted as $1s^2 2s^1$. As can be seen from the first column of the periodic table, a lone s^1 electron in the outermost principal orbital is the characteristic of all *alkali metals*.

Beryllium with $Z = 4$ has an electronic configuration of $1s^2 2s^2$. After beryllium, the electrons start to fill the states of the second principal orbital. Two linear combinations (addition and subtraction) of the electron density clouds corresponding to $m_l = +1$ and -1 yield the p_x and p_y orbitals. The p_z orbital results when $m_l = 0$. In contrast to the s orbital, *the probability cloud of a p orbital is not spherically symmetric*. Each of p_x , p_y and p_z orbitals has the maximum electron density along one of the three coordinate directions, as shown in Fig. 4.2. When bonds are formed by electrons of the p orbitals, this directional nature of the orbital plays an important role in determining the bond angles. The electron configuration of the elements from lithium to neon is summarized in Table 4.1.

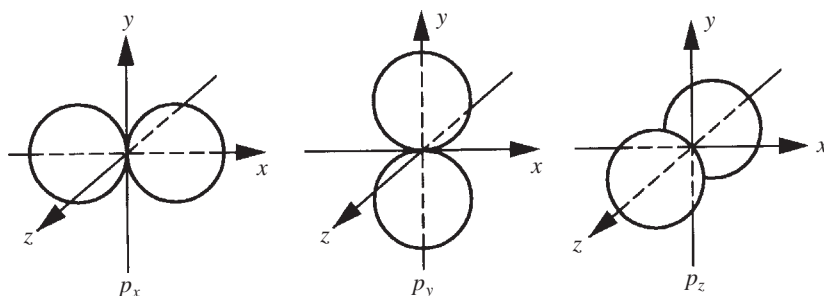


Fig. 4.2 The electron density clouds of p_x , p_y and p_z orbitals are concentrated each along one of the three coordinate axes.

Hund's rule states that, in order to reduce the electron–electron repulsive energy, the number of electrons of the same spin in p , d or f states should be maximum. The filling order in the p orbital shown in Table 4.1 is in accord with

TABLE 4.1
Elements in the Second Row of the Periodic Table

Element	Symbol	Atomic number	Electronic configuration	Electron spin distribution in p states		
				p_x	p_y	p_z
Lithium	Li	3	$1s^2 2s^1$	—	—	—
Beryllium	Be	4	$1s^2 2s^2$	—	—	—
Boron	B	5	$1s^2 2s^2 2p^1$	↑	—	—
Carbon	C	6	$1s^2 2s^2 2p^2$	↑	↑	—
Nitrogen	N	7	$1s^2 2s^2 2p^3$	↑	↑	↑
Oxygen	O	8	$1s^2 2s^2 2p^4$	↑↓	↑	↑
Fluorine	F	9	$1s^2 2s^2 2p^5$	↑↓	↑↓	↑
Neon	Ne	10	$1s^2 2s^2 2p^6$	↑↓	↑↓	↑↓

this rule. Nitrogen, for example, has three electrons in the p orbital. They are distributed one each in the p_x , p_y and p_z orbitals so that all three can have the same spin.

The halogen fluorine is only one electron short for the full complement of p electrons—a characteristic common to all the halogens. When all the electrons are present in the p orbital, the sum of the electron probability clouds along the three coordinate axes results in spherical symmetry. Neon has all the six p electrons and has a spherically symmetric cloud around it. This configuration is called *the inert gas configuration*. All the elements at the end of a row in the periodic table are inert gases.

The *third row* of the periodic table starts (as in the second row) with an alkali metal (sodium) having a s^1 configuration in the outermost orbital. The

complete electron configuration of sodium is $1s^22s^22p^63s^1$. The row ends with the inert gas argon, with configuration $1s^22s^22p^63s^23p^6$. The number of quantum states in the third principal orbital is 18 since, for $n = 3$, we can have

$$\left. \begin{array}{l} l = 0, m_l = 0, m_s = \pm 1/2 \\ l = 1, m_l = -1, m_s = \pm 1/2 \\ m_l = 0, m_s = \pm 1/2 \\ m_l = +1, m_s = \pm 1/2 \end{array} \right\} 3s$$

$$\left. \begin{array}{l} l = 1, m_l = -1, m_s = \pm 1/2 \\ m_l = 0, m_s = \pm 1/2 \\ m_l = +1, m_s = \pm 1/2 \end{array} \right\} 3p$$

$$\left. \begin{array}{l} l = 2, m_l = -2, m_s = \pm 1/2 \\ m_l = -1, m_s = \pm 1/2 \\ m_l = 0, m_s = \pm 1/2 \\ m_l = +1, m_s = \pm 1/2 \\ m_l = +2, m_s = \pm 1/2 \end{array} \right\} 3d$$

However, the $3d$ states corresponding to $l = 2$ remain vacant till the fourth row is reached in accordance with the order of increasing energy of the orbitals. Therefore, the third row of the periodic table has only eight elements.

The *fourth row* starts with potassium of configuration $1s^22s^22p^63s^23p^64s^1$. Next comes calcium. After calcium, the first transition series begins, where the $3d$ orbitals begin to get filled. The d orbitals are directional in nature like p orbitals and can hold a maximum of 10 electrons. The order of filling of $3d$ orbitals is summarized in Table 4.2. When the d shell is exactly half-filled as in chromium, we have five electrons of the same spin in the d shell.

TABLE 4.2
Elements of the First Transition Series

Element	Symbol	Atomic number	Outer electron configuration	Electron spin distribution in d states				
				d_1	d_2	d_3	d_4	d_5
Scandium	Sc	21	$3s^23p^63d^1 4s^2$	↑	—	—	—	—
Titanium	Ti	22	$3s^23p^63d^2 4s^2$	↑	↑	—	—	—
Vanadium	V	23	$3s^23p^63d^3 4s^2$	↑	↑	↑	—	—
Chromium	Cr	24	$3s^23p^63d^5 4s^1$	↑	↑	↑	↑	↑
Manganese	Mn	25	$3s^23p^63d^5 4s^2$	↑	↑	↑	↑	↑
Iron	Fe	26	$3s^23p^63d^6 4s^2$	↑↓	↑	↑	↑	↑
Cobalt	Co	27	$3s^23p^63d^7 4s^2$	↑↓	↑↓	↑	↑	↑
Nickel	Ni	28	$3s^23p^63d^8 4s^2$	↑↓	↑↓	↑↓	↑	↑
Copper	Cu	29	$3s^23p^63d^{10} 4s^1$	↑↓	↑↓	↑↓	↑↓	↑↓
Zinc	Zn	30	$3s^23p^63d^{10} 4s^2$	↑↓	↑↓	↑↓	↑↓	↑↓

When all the ten electrons are filled, the d orbital acquires spherical symmetry and we move out of the transition series. The copper atom with the outer configuration $3d^{10}4s^1$ has a spherical electron probability cloud around the nucleus. Next to copper, zinc has the outer configuration $3d^{10}4s^2$. The $4p$ orbitals begin to get filled from gallium onwards up to the inert gas krypton (see the Periodic Table).

The *fifth row* contains the second series of transition elements and is similar to the fourth row, in the order of filling. The $4f$ and the $5f$ orbitals are filled in the sixth and the seventh rows of the periodic table. There are 14 f -orbitals corresponding to the value of $l = 3$ and $m_l = -3, -2, -1, 0, +1, +2$, and $+3$. Elements with partially filled f orbitals are called the *rare earth elements*. The sixth row has 32 elements in it, while the seventh row goes up to lawrencium ($Z = 103$). In recent years, some elements with $Z > 103$ have been discovered.

4.3 Ionization Potential, Electron Affinity and Electronegativity

We have already noted that electrons in the orbitals around the nucleus can absorb energy and get promoted to higher energy levels. If sufficient energy is supplied, an electron in the outer orbital can break away completely from the atom and become free. The energy required to remove an electron in this manner is known as the *ionization potential*. The energy to remove the outermost electron, which is weakly bound to the atom, is called the *first ionization potential*.

Figure 4.3 gives the first ionization potentials of the elements. The potential is the least for the first atom in a row of the periodic table and increases as we

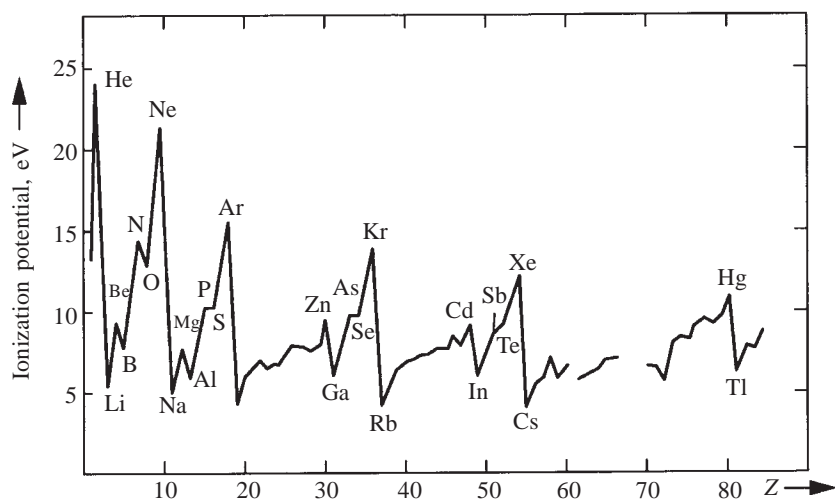


Fig. 4.3 First ionization potentials of the elements.

go to the right along the row. The alkali metals have a lone s electron in their outermost orbital, which can be removed with relative ease. The inert gases, on the other hand, have the full complement of s and p electrons in their outermost principal orbital. Removing an electron from the stable inert gas configuration requires a relatively large expenditure of energy. Further, as we go down a column of the periodic table, the outermost electrons are less and less tightly bound to the nucleus. Correspondingly, among the alkali metals, the first ionization potential is highest for lithium at the top of the column and lowest for cesium at the bottom. Among the inert gases, the potential is highest for helium and lowest for xenon.

When an electron is removed from the neutral atom, there is a decrease in the mutual repulsion between the orbital electrons. They can approach one another more closely and are therefore attracted more strongly to the nucleus, resulting in the shrinking of all the orbitals. As a result, the energy required to remove the second and successive electrons becomes increasingly greater.

Example 4.2 Give the electronic configuration of a neutral iron atom, a ferrous ion and a ferric ion. Compare their sizes.

Solution

Species	Electronic configuration	Radius, Å
Neutral iron Fe	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^6 4s^2$	1.24
Ferrous ion Fe^{2+}	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^6$	0.83
Ferric ion Fe^{3+}	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^5$	0.67

The size of the atom decreases considerably as more and more electrons are removed from the outer orbital.

Consider a system of a neutral atom and an extra electron. The work done by this system, when the extra electron is attracted from infinity to the outer orbital of the neutral atom, is known as the *electron affinity* of the atom. The electron affinities of some elements are shown in Table 4.3. The stable configuration of the inert gases has no affinity for an extra electron. The halogens, which are just one electron short to achieve the stable inert gas configuration, have the largest electron affinities. When an extra electron is added to a neutral atom, there is a weakening of the attraction of the electrons to the nucleus, resulting in an expansion of the electron orbitals and an increase in the size of the atom.

TABLE 4.3
Electron Affinities of Some Elements
(Values are given in eV and kJ mol⁻¹)

For one electron							
H							He
0.7							0
68							0
Li	Be	B	C	N	O	F	Ne
0.54	0	0.54	1.13	0.2	1.48	3.62	0
52	0	52	109	19	143	349	0
Na	Mg	Al	Si	P	S	Cl	Ar
0.74	0	0.4	1.90	0.80	2.07	3.82	0
71	0	39	183	77	200	369	0
						Br	Kr
						3.54	0
						342	0
						I	Xe
						3.24	0
						313	0

The tendency of an atom to attract electrons to itself during the formation of bonds with other atoms is measured by the *electronegativity* of that atom. This is not the same as the electron affinity, where the tendency of an atom to attract an isolated electron is measured. Consider, for example, the bond formation between hydrogen and fluorine by the sharing of their outer electrons. Fluorine has a greater tendency to attract the bonding electrons to itself than hydrogen. Fluorine has thus a larger electronegativity than hydrogen. Pauling has worked out empirically the electronegativities of elements, Table 4.4. As seen from the table, the halogens have the largest electronegativities and the alkali metals have the smallest. Qualitatively, the nonmetals such as the halogens are said to be electronegative. A value of 2 on the electronegativity scale can be taken as the approximate dividing line between metals and nonmetals.

TABLE 4.4
Pauling's Electronegativities of Elements

H																			
2.1																			
Li	Be											B	C	N	O	F			
1.0	1.5											2.0	2.5	3.0	3.5	4.0			
Na	Mg											Al	Si	P	S	Cl			
0.9	1.2											1.5	1.8	2.1	2.5	3.0			
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br			
0.8	1.0	1.3	1.5	1.6	1.6	1.5	1.8	1.8	1.8	1.9	1.6	1.6	1.8	2.0	2.4	2.8			
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I			
0.8	1.0	1.2	1.4	1.6	1.8	1.9	2.2	2.2	2.2	1.9	1.7	1.7	1.8	1.9	2.1	2.5			
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At			
0.7	0.9	1.1	1.3	1.5	1.7	1.9	2.2	2.2	2.2	2.4	1.9	1.8	1.8	1.9	2.0	2.2			

CHEMICAL BONDING

4.4 Bond Energy, Bond Type and Bond Length

The forces between two atoms (or ions) as a function of their distance of separation r are schematically shown in Fig. 4.4. When the distance of

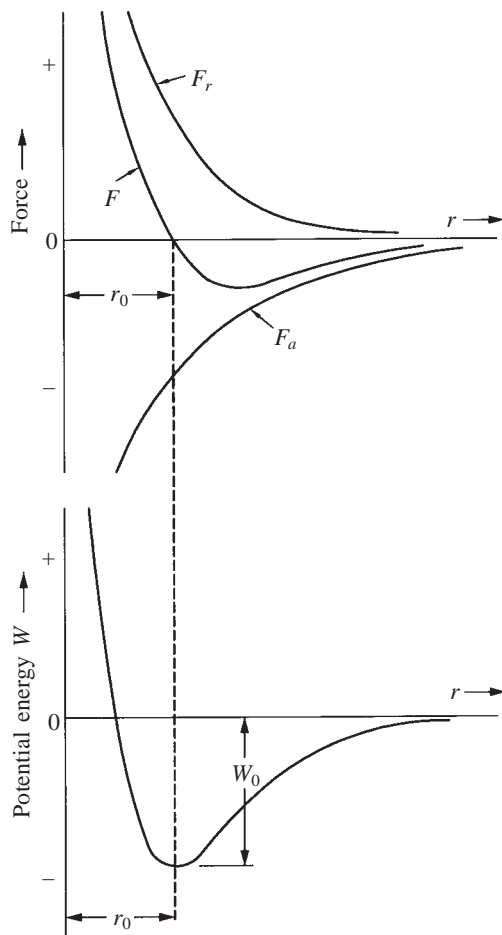


Fig. 4.4 Interatomic forces and potential energy in a system of two atoms, as a function of their distance of separation r .

separation is large, the significant force is the attractive (negative) force, F_a . At closer distances of approach, a repulsive (positive) force, F_r , also becomes significant and increases rapidly with decreasing distance of separation. When the distance of separation is r_0 , the attractive and the repulsive forces exactly balance each other and the net force is zero. This distance corresponds to stable equilibrium with a minimum in potential energy W . The magnitude of the minimum energy W_0 is called the *bond energy*.

Bond energy should be expressed in kilojoules per mole of bonds (kJ mol^{-1}). Other units such as electron volt per bond and kilocalories per mole of bonds are also used. It is useful to remember that 1 eV/bond is approximately equal to 100 kJ mol^{-1} and that 1 kilocalorie is equal to 4.18 kilojoules . The energy per mole of bonds is not necessarily equal to the enthalpy of atomization which is the energy required to convert one mole of atoms in a solid into its component atoms in the gaseous state. This is so because each atom in the solid may form bonds with a number of neighbours, and all these bonds are broken during vapourization. If the bonds are discrete, the enthalpy of atomization equals $z/2$ times the energy per mole of bonds, where z is the number of nearest bonding neighbours of an atom. The factor 2 arises as each bond between a pair of neighbours is counted twice.

According to strength, chemical bonds can be grouped into primary and secondary bonds. *Primary bonds* have bond energies in the range $100\text{--}1000 \text{ kJ mol}^{-1}$ ($1\text{--}10 \text{ eV/bond}$). Covalent, metallic and ionic bonds are all primary bonds. Among these, covalent and ionic bonds are generally stronger than metallic bonds. *Secondary bonds* have energies in the range $1\text{--}50 \text{ kJ mol}^{-1}$ ($0.01\text{--}0.5 \text{ eV/bond}$), one or two orders of magnitude smaller than those of primary bonds. Examples of secondary bonds are van der Waals bonds and the hydrogen bond. Generally, van der Waals bonds are very weak.

Few materials have pure bonds of one type or the other. Occurrence of mixed bonds in materials is the rule rather than the exception. In fact, empirical rules exist to apportion a fraction of a bond to a particular type. Notwithstanding this, it is useful to classify materials according to the bond type that is *dominant* in a given material. This classification helps in predicting the approximate properties and behaviour of a material under a given set of conditions even if the actual behaviour may sometimes not agree with this prediction.

The *length of a bond* (r_0 in Fig 4.4) is defined as the centre-to-centre distance of the bonding atoms. Strong bonds pull the bonding atoms closer together and so have smaller bond lengths as compared to weak bonds. Primary bonds have lengths in the range $1\text{--}2 \text{ \AA}$ ($0.1\text{--}0.2 \text{ nm}$). Secondary bond lengths are larger, in the range $2\text{--}5 \text{ \AA}$ ($0.2\text{--}0.5 \text{ nm}$).

The length of a bond can be used to define atomic or ionic diameters. When bonding is between two neighbouring atoms of the same kind, the atomic diameter is simply equal to the bond length, see Fig. 4.5a. For example, the equilibrium distance between the atomic centres in the diatomic chlorine molecule is 1.81 \AA , which is also the diameter of the chlorine atom. To indicate the character of the bond, this can be called the covalent diameter of chlorine. The equilibrium distance between two nearest bonding copper atoms in a copper crystal is 2.56 \AA , which is also the (metallic) diameter of copper. Some ambiguity in this definition arises if the element in question exhibits different crystal forms. For example, the diameter of the iron atom is 2.48 \AA when it is surrounded by eight neighbours in the BCC crystal and 2.54 \AA when it has 12 nearest neighbours in the FCC crystal. When two bonding atoms are of different types, as in the ionic bonding, the bond length is equal to the sum of their radii, $r_c + r_a$, as illustrated in Fig. 4.5b.

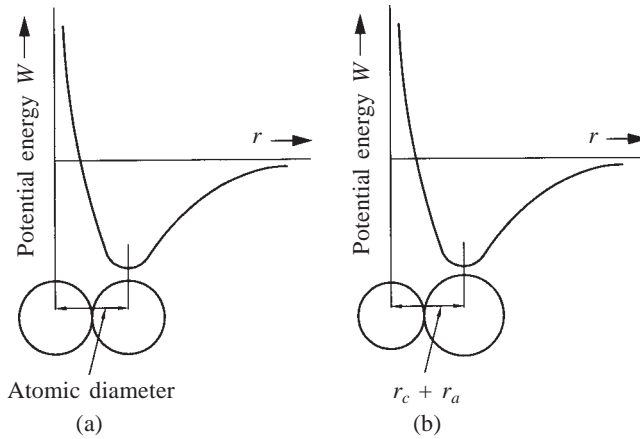


Fig. 4.5 The relationship between atomic (or ionic) size and the bond length.

The equilibrium distance of separation shown in Fig. 4.4 applies to 0 K, where there is no thermal energy. At higher temperatures, under the influence of thermal energy, atoms vibrate about their mean positions, the amplitude of vibrations increasing with increasing temperature. As can be seen from Fig. 4.6, at temperature T_1 , the amplitude is a_1b_1 and at T_2 ($T_2 > T_1$) the amplitude is a_2b_2 . The corresponding mean spacings between the atoms are given by r'_0 and r''_0 . As the repulsive force is short range in nature as compared to the attractive force, the potential energy curve is steeper on the left side of r_0 than on the right side. Therefore, $r''_0 > r'_0 > r_0$ as shown in Fig. 4.6. That is, the mean bond length increases on heating. In other words, the material exhibits *thermal expansion*.

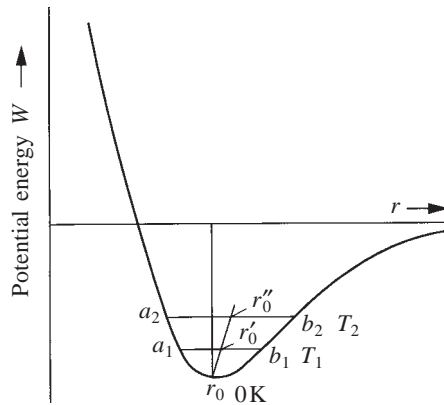


Fig. 4.6 Effect of temperature on the mean spacing between atoms forming a chemical bond.

4.5 Ionic Bonding

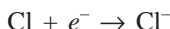
Ionic bonding forms between two oppositely-charged ions which are produced by the transfer of electrons from one atom to another. The steps in the formation of an ionic bond between sodium and chlorine can be visualized as follows:

4.5.1 Production of Ions of Opposite Sign

The element sodium has an atomic number $Z = 11$ and an electronic configuration $1s^2 2s^2 2p^6 3s^1$. Being an alkali metal, the first ionization potential for sodium is relatively small, 496 kJ mol^{-1} (5.1 eV/atom). The outermost electron is removed by supplying this much energy:



The released electron is then available to occupy the only vacant state in the chlorine atom to produce a negatively charged ion:



The electron affinity of chlorine is 369 kJ mol^{-1} (3.8 eV/atom). Thus, the electron transfer from sodium to chlorine produces two oppositely charged univalent ions, with a net increase in the potential energy $\Delta E = 496 - 369 = 127 \text{ kJ mol}^{-1}$ (1.3 eV/pair). This increase is more than compensated during bonding by the energy decrease due to electrostatic attraction between the two oppositely charged ions.

4.5.2 The Coulomb Attraction

The electron transfer results in the inert gas configuration around both the nuclei. This configuration has spherical symmetry of the electron probability cloud. So, the bonding force between the ions is the same in all directions. The ionic bond is said to be *nondirectional*. To start with, the ions can be considered to be point charges. From Coulomb's law of electrostatics, as two opposite charges are brought together, the attractive force between them increases in magnitude inversely as the square of the distance of separation r . So, the potential energy decreases inversely as the distance. The potential energy is equal to $-A'z_1z_2e^2/r$, where z_1 and z_2 are the valency of the two ions, e is the electronic charge and A' is a conversion factor.

4.5.3 The Short Range Repulsion

The assumption that the ions are point charges will be valid only when the electron clouds surrounding the two nuclei do not overlap. The Pauli exclusion principle will come into the picture if there is an appreciable overlap between the two electron clouds. Not more than one electron can occupy a given quantum state. This gives rise to a repulsive force between the ions, which is short range in nature. When the electron clouds start to overlap, the repulsive force increases sharply with decreasing distance. The repulsive energy is taken to vary as the reciprocal of some power m ($m > 1$) of the distance of separation r . It is represented by B/r^m , where B and m are empirically determined constants. The value of m varies between 9 and 15.

Summing up the three steps, the potential energy W of the system of the bond forming atoms is given by

$$W = \frac{-A'z_1z_2e^2}{r} + \frac{B}{r^m} + \Delta E \quad (4.4)$$

The minimum in the energy function can be obtained by setting $dW/dr = 0$, to yield the bond energy W_0 and the equilibrium bond length r_0 . The strength of an ionic bond will depend on the relative magnitudes of the three energy terms on the right side of Eq. (4.4). Considering ΔE first, this must be as small as possible. Electropositive elements such as the alkali metals have small ionization potentials. Electronegative elements such as halogens have large electron affinities. Hence ionic bonds form most readily between electropositive and electronegative elements. ΔE is almost zero for electron transfer from a cesium atom to a chlorine atom. Consequently, these two elements form a bond that is dominantly ionic in character. The Coulomb term is dependent on the charges of the ions forming the bond. Ionic crystals with multivalent ions have generally stronger bonds and hence higher melting points than crystals with univalent ions. Thus, MgO, BeO and Al_2O_3 are refractory oxides with melting points above 2000°C . MgO (magnesia) has important applications as a refractory in the steel making industry. Al_2O_3 (alumina) crucibles can hold almost all common metals in the molten condition.

As ionic bonds are nondirectional, a cation in a crystal tends to surround itself by as many anions as possible and vice versa. The Coulomb term for a crystal is then dependent on the mutual attractive as well as repulsive interaction of a given ion with all the other ions in the crystal, e.g., attractive interaction with the nearest neighbours, repulsive interaction with the next-nearest neighbours (which are of the same sign) and so on. This is an interesting problem in the spatial array of the ions and was first solved by Madelung. The Coulomb energy term for a pair of isolated ions is simply multiplied by a constant to obtain the energy of a pair of ions in the crystal. This constant called the *Madelung constant* is 1.748 for a NaCl-type crystal and 1.763 for a CsCl-type crystal.

The enthalpies of atomization of some typical ionic crystals are shown in Table 4.5.

TABLE 4.5
Enthalpy of Atomization of Some Ionic Crystals

Crystal	Enthalpy of atomization, kJ mol^{-1} (of compound)	Melting point, $^\circ\text{C}$	Boiling point, $^\circ\text{C}$
LiF	864	850	1680
NaCl	640	801	1445
KCl	669	770	1415
CsCl	649	605	1300
MgO	1000	2850	3260
BeO	1170	2550	—
CaF_2	1548	1420	2510
Al_2O_3	3060	2050	—

4.6 Covalent Bonding

Covalent bonding occurs by the *sharing of electrons* between neighbouring atoms. This is in contrast to the transfer of electrons from one atom to another in the ionic bonding. For sharing with a net decrease in potential energy, good overlap of the orbitals, which will bring the shared electrons close to both the nuclei, is necessary. This occurs readily when there are vacant electron states in the outermost orbital of the bonding atoms. When the overlapping orbitals are directionally oriented and not spherically symmetric, good overlapping and substantial decrease in the potential energy can occur. This also gives *directionality* to the covalent bond.

Consider first the formation of a hydrogen molecule. When two hydrogen atoms are very far apart, they do not interact, and the lone electrons of the atoms stay in their respective $1s$ ground states. When the atoms come closer, the electron probability clouds of the $1s$ states overlap. As the $1s$ orbitals can have two electrons of opposite spin, the sharing of electrons between the two atoms takes place, without having to promote the electrons to higher energy levels. Both the electrons are close to both the nuclei and, in fact, spend much of the time in between the two nuclei. The hydrogen molecule has a bond energy 436 kJ mol^{-1} and a bond length of 0.74 \AA . The next element helium cannot form a covalent bond as this requires promotion of the $1s$ electrons to the second principal orbital.

Sharing of electrons and the formation of covalent bonds readily occurs between atoms which have unfilled p orbitals. The p orbitals are directional in nature and hence permit efficient overlapping of the orbitals in the direction of the maximum electron probability density. Fluorine with $Z = 9$ and an electronic configuration $1s^2 2s^2 2p^5$ has a vacant state in the p_z orbital. Two fluorine atoms can come together such that the half-filled p_z orbitals overlap. As there are no other half-filled orbitals, each fluorine atom forms only one bond. In the elemental state, fluorine forms the diatomic molecule F_2 with a bond energy of 154 kJ mol^{-1} and a bond length of 1.42 \AA .

In oxygen to the left of fluorine, the p_y and p_z orbitals have one vacant state each. Each of the unpaired electrons in these orbitals can share an electron with another atom. In the water molecule, the unpaired $1s$ electrons of two hydrogen atoms pair up with the two unpaired $2p$ electrons of oxygen. If the overlap occurred without any distortion of the p orbitals, the angle between the bonds would be the same 90° as between the unbonded p orbitals. However, the observed H–O–H bond angle in the water molecule is 104° . The formation of the bond here is more appropriately visualized to occur between *hybridized orbitals* produced by the interaction between the $2s$ orbitals and the $2p$ orbitals of the oxygen atom. The hybridized orbitals, known as the (sp^3) orbitals, are four in number and can hold a total of eight electrons. Their mutual orientation is ideally the one between the four lines joining the centre to the four corners of a regular tetrahedron. (A regular tetrahedron is a solid figure made up of four faces, all of which are equal and equilateral triangles.) The interbond angle corresponding to this orientation is 109.5° . In the water molecule, two of the

hybridized orbitals of oxygen are occupied by four electrons of oxygen and hence do not take part in bonding. In the other two orbitals, electrons are shared between two hydrogen atoms. By a similar bonding mechanism, the other elements of the sixth column form two bonds. That is, each atom is bonded to two neighbours. The S–S–S bond in sulphur has a bond angle of 107° . The Se–Se–Se bond in selenium and the Te–Te–Te bond in tellurium have a bond angle of 104° .

The above are examples of an *end-to-end overlap* of p orbitals giving rise to what is known as a σ (*sigma*) bond. When there is *lateral overlap* of p orbitals, π (*pi*) bonds are said to form. Double and triple bonds are examples of π bonds. Figure 4.7 illustrates the difference between the end-to-end overlap and the lateral overlap of p orbitals. In the diatomic oxygen molecule, one double bond holds together the two oxygen atoms, with a bond energy of 494 kJ mol^{-1} and a bond length of 1.21 \AA .

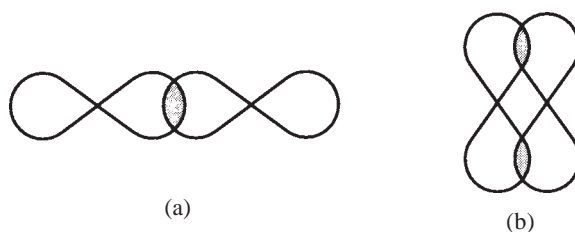


Fig. 4.7 (a) End-to-end overlap gives rise to a σ bond. (b) Lateral overlap gives rise to a π bond.

A fifth column element has all three p electrons unpaired. Hence an atom here forms three covalent bonds. The bond angles reflect the fact that the three p orbitals have a mutually perpendicular orientation. Some examples are given in Table 4.6.

TABLE 4.6
Bonding Characteristics of Fifth Column Elements

Element	Atomic number	Bond	Bond angle	Covalent radius, \AA	Bond energy, kJ mol^{-1}
Phosphorus	15	P–P–P	99°	1.10	214
Arsenic	33	As–As–As	97°	1.25	134
Antimony	51	Sb–Sb–Sb	96°	1.45	126
Bismuth	83	Bi–Bi–Bi	94°	1.56	105

These bond angles are closer to the 90° angle between unbonded p orbitals, rather than to the tetrahedral angle of 109.5° , which is characteristic of hybridized orbitals. With increasing atomic number, the bond length increases and the bond strength decreases. As we go down the column, the increasing shielding effect of the electrons of the inner orbitals binds the outer electrons

less and less tightly to the nucleus. This explains the decrease in the bond strength as the atomic number increases. Nitrogen in this column forms a diatomic molecule with a triple bond between two nitrogen atoms, which has a bond energy of 942 kJ mol^{-1} and a bond length of 1.10 \AA .

The *diamond form* of carbon and other fourth column elements share all the four hybridized orbitals with four neighbours. *The four bonds formed are of equal strength and bear the ideal tetrahedral angle of 109.5° to one another.* They form crystals with a three-dimensional network of covalent bonds. The bonding characteristics are given in Table 4.7.

TABLE 4.7
Bonding Characteristics of Fourth Column Elements

Material	Atomic number	Bond energy, kJ mol^{-1}	Bond length, \AA	Melting point, $^\circ\text{C}$	Boiling point, $^\circ\text{C}$
Diamond C	6	347	1.54	3550 (gr)	4000 (gr)
Silicon Si	14	176	2.36	1410	2680
Germanium Ge	32	159	2.44	937	2830
Gray tin Sn	50	146	3.02	232	2270
Silicon carbide SiC	—	308	1.88	~2500	—

The first four are elements of the fourth column with increasing atomic number. As with the fifth column elements, they form bonds of decreasing strength and increasing bond length.

Example 4.3 The enthalpy of atomization of diamond is 713 kJ mol^{-1} . The listed bond energy is 347 kJ mol^{-1} . How do you reconcile this difference?

Solution In diamond, each carbon atom forms four discrete bonds. As two atoms are required to form a bond, the effective number of bonds per atom is $4/2 = 2$. So, when one mole of diamond is converted into carbon atoms in the gaseous state, 2 moles of C–C bonds are broken. $347 \times 2 = 694 \text{ kJ mol}^{-1}$ should be the enthalpy of atomization of diamond. It is close to the given value.

The bond energy of single, double and triple bonds increases in that order, with a corresponding decrease in the bond length. This is illustrated in Table 4.8 for carbon.

TABLE 4.8
Single and Multiple Bonds of Carbon

Nature of bond	Bond energy, kJ mol^{-1}	Bond length, \AA
Single C–C	347	1.54
Double C=C	614	1.33
Triple C \equiv C	811	1.20

4.7 Metallic Bonding

The elements to the left of the fourth column in the periodic table exhibit metallic characteristics. The sharing of electrons between neighbouring atoms now becomes *delocalised* as there are not enough electrons to produce the inert gas configuration around each atom. The metallic sharing changes with time and the bonding electrons resonate between different atoms. The metallic state can be visualized as an array of positive ions, with a common pool of electrons to which all the metal atoms have contributed their outer electrons. This common pool is called the *free electron cloud* or the free electron gas. These electrons have freedom to move anywhere within the crystal and act like an all-pervasive, mobile glue holding the ion cores together. This is in sharp contrast to the electrons in covalent bonding, which are localised, bind just two neighbouring atoms, and stay with them. This freedom makes the metallic bonds *nondirectional*. Note that the ion cores have spherical electron density clouds around them. The bond energies of common metals are listed in Table 4.9.

TABLE 4.9
Bond Energies of Metals

Metal		Bond energy, kJ mol ⁻¹ (of bonds)	No. of bonding neighbours of an atom	Melting point, °C	Boiling point, °C
Copper	Cu	56.4	12	1083	2595
Silver	Ag	47.5	12	961	2210
Gold	Au	60.0	12	1063	2970
Aluminium	Al	54.0	12	660	2450
Nickel	Ni	71.6	12	1453	2730
Platinum	Pt	94.3	12	1769	4530
Lead	Pb	32.5	12	327	1725
Cobalt	Co	70.8	12	1495	2900
Magnesium	Mg	24.6	12	650	1107
Cadmium	Cd	18.6	12	321	765
Zinc	Zn	21.9	12	420	906
Sodium	Na	27.0	8	98	892
Tungsten	W	212.3	8	3410	5930
Iron	Fe	104.0	8	1535	3000
Chromium	Cr	99.3	8	1875	2665
Molybdenum	Mo	164.6	8	2610	5560
Vanadium	V	128.5	8	1900	3400
Niobium	Nb	180.8	8	2468	4927

Bond lengths can be derived from the data in the table on the back inside cover.

4.8 Secondary Bonding

In many molecules, where hydrogen takes part in the covalent bonding, the centres of the positive and the negative charges do not coincide. Consider the example of the water molecule. The electronegativity of oxygen is 3.5 and that of hydrogen is 2.1. Therefore, the oxygen atom pulls the bonding electrons to itself more strongly than hydrogen does. This results in a net negative charge at the oxygen end and a net positive charge at the hydrogen end of the molecule. Due to this imbalance in electrical charge, the water molecule possesses a *permanent dipole moment*. The bond that is formed between water molecules due to attraction between the positively-charged hydrogen end of a molecule and the negatively-charged oxygen end of another molecule is called *the hydrogen bond*, see Fig. 4.8.

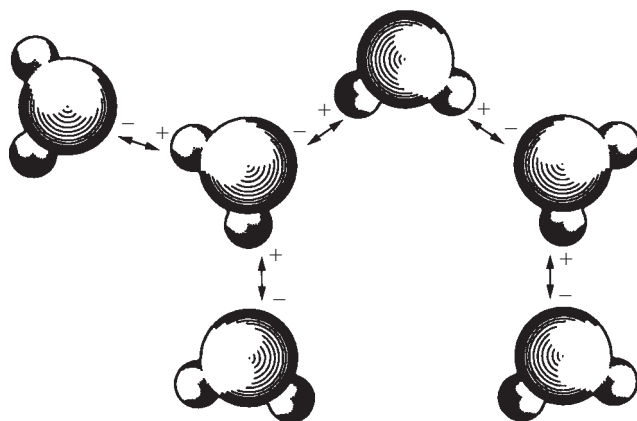


Fig. 4.8 Hydrogen bond between water molecules.

Hydrogen bonds are evidently directional. The bond between water molecules is strong enough to persist in the liquid state. It is responsible for the unusual properties of ice and water. For example, the relatively open network of hydrogen bonds in ice as shown in Fig. 4.8 collapses to a more closely packed liquid, accounting for the anomalous increase in density on melting. The hydrogen bond energies and the permanent dipole moments of some common molecules are shown in Table 4.10.

TABLE 4.10

Hydrogen Bond Energies and Dipole Moments of Some Molecules

Molecule	Dipole moment		Bond energy, kJ mol^{-1}	Melting point, $^{\circ}\text{C}$	Boiling point, $^{\circ}\text{C}$
	C m	Debye			
Water H_2O	6.2×10^{-30}	1.85	20.5	0	100
Ammonia NH_3	4.9×10^{-30}	1.48	7.8	-78	-33
HF	6.7×10^{-30}	2.00	31.5	-83	20

Example 4.4 From the enthalpy of fusion of ice, estimate the fraction of hydrogen bonds that are broken when ice melts.

Solution The enthalpy of fusion of ice = 6.02 kJ mol^{-1} . There are two moles of hydrogen bonds per mole of H_2O in ice. From Table 4.10, the hydrogen bond energy = 20.5 kJ mol^{-1} . Assuming that all the heat absorbed during melting goes into breaking the bonds, the fraction of bonds broken is given by

$$6.02/(20.5 \times 2) = 0.15$$

As bonds are continuously broken and remade in a liquid, this should be considered as a time-averaged value.

Inert gas atoms have spherically symmetric electron probability clouds around them and, therefore, have no permanent dipole moments. Yet, inert gases form solid crystals at sufficiently low temperatures. The bonding in such solids is called the *van der Waals bonding*. It is the result of *momentary fluctuations* in the charge distribution around an atom. Even though the time averaged electron probability distribution is spherically symmetric, the electronic charge at any instant of time is concentrated locally, resulting in a weak fluctuating dipole within the atom. The electric field of this imbalance can induce a dipole moment in a neighbouring atom, in such a way as to attract it. The dipole in the second atom can in turn induce a dipole in a third atom in order to attract it, and so on. This dipole-induced-dipole attraction is *nondirectional* in nature. When the atoms come closer, a repulsion arises, which can be explained (as in other bonds) on the basis of the Pauli exclusion principle. The bond energies and the bond lengths of some inert-gas crystals formed by van der Waals attraction are given in Table 4.11. The very low bond energies are associated with correspondingly large bond lengths.

TABLE 4.11
Bonding Characteristics of Some Inert-Gas Crystals

Crystal	Symbol	No. of bonding neighbours of an atom	Bond energy, kJ mol^{-1}	Bond length, \AA	Melting point, $^{\circ}\text{C}$	Boiling point, $^{\circ}\text{C}$
Helium	He	12	0.29	3.57	-270	-269
Neon	Ne	12	0.30	3.15	-249	-246
Argon	Ar	12	1.09	3.76	-189	-186
Krypton	Kr	12	1.50	3.99	-157	-152
Xenon	Xe	12	2.10	4.34	-112	-108

4.9 Variation in Bonding Character and Properties

We will examine some generalizations between properties and the bonding character of a material. The solid state can be visualized as atoms vibrating about their mean positions on fixed atomic sites. In the liquid state, the atoms have also translational freedom and can slide past one another. The bonds between atoms in the liquid are continuously broken and remade. In the gaseous state, the bonds are totally broken. The thermal energy of atoms must be sufficient to achieve these disruptions in bonding. The higher the bond strength, the more will be the thermal energy required to break the bonds. Correspondingly, *strongly bonded materials tend to have high melting and boiling temperatures*. Among the primary bonds, covalent and ionic bonds are generally stronger than metallic bonds. Hence, covalent and ionic solids have high melting and boiling points. The general trend of correlation between bond strength and the melting and boiling temperatures is already shown in several tables of this chapter.

When a solid consists of molecules held together by secondary bonds, the melting and boiling points of the solid reflect only the strength of the secondary bonds *between* the molecules, and not the strength of the primary bonds *within* the molecule. Here, the molten state and the gaseous state are to be visualized as consisting of units of molecules and not individual atoms of the molecule. The silica (SiO_2) crystal, which has a three-dimensional network of Si–O bonds without any secondary bonding, has a melting point of 1723°C . During melting, the Si–O bonds are broken. In sharp contrast to this, methane (CH_4) has C–H bonds in the molecule. The molecules are held together by van der Waals bonds in the solid which melts at -182°C . Here, the C–H bond strength of 413 kJ mol^{-1} , which is even larger than the Si–O bond strength of 375 kJ mol^{-1} , has no correlation with the melting point of methane. Rather, the van der Waals bond strength of 1.36 kJ mol^{-1} is to be associated with the low melting temperature.

The thermal and electrical conductivities of a solid are, to a large extent, dependent on the presence of *free electrons* in the solid. In ionic solids, the electron transfer produces the stable inert gas configuration around both the cations and the anions. Hence, there are no free electrons in ionic solids. Likewise, covalent bonding produces the inert gas configuration around atoms sharing the electrons, with the result that there are no free electrons here either. Hence, *typical ionic and covalent materials are good thermal and electrical insulators*. Solids which have secondary bonds such as van der Waals bonds (wholly or in addition to ionic or covalent bonds) are also good insulators. In contrast, metals have free electrons and are therefore good conductors of heat and electricity. The best known conductors of heat and electricity are copper, silver and gold.

We noted that the *thermal expansion* of materials arises from the asymmetry of the potential energy versus distance curve (refer Fig. 4.6). Deep potential wells are more symmetrical about the equilibrium position r_0 than shallow

potential wells. So, the thermal expansion at a given temperature tends to be less for strongly bonded materials than for weakly bonded materials. The thermal expansion coefficients of a number of materials at room temperature are given in Appendix I.

The *mechanical properties* of solids are dependent on the strength of the bonds as well as the directional nature of bonding. Solids with strong and directional bonds tend to be brittle. For example, covalently bonded diamond is very hard and brittle. As metallic bonds are relatively weak and nondirectional, metals are soft, ductile and malleable. They can change their shape permanently without breaking. Ionic solids fall in between covalent and metallic solids in that they may exhibit a very limited amount of ductility.

In any one row of the periodic table, as we go from left to the right, the metallic character of the bond decreases and the covalent character increases. Metallic bonds being weaker than covalent bonds, this transition is reflected in the increasing bond energy and the decreasing bond length from left to the right, as shown in Table 4.12 for the third row of the periodic table. Similarly, the transition from covalent to metallic character, as we go from top to the bottom of a column, is seen in the decreasing bond energy in Table 4.7.

TABLE 4.12
Properties of Elements of the Third Row

Element	Na	Mg	Al	Si	P	S	Cl
Bond energy, kJ mol ⁻¹	27	25	54	176	214	243	242
Bond length, Å	3.72	3.18	2.86	2.36	2.20	2.08	1.81
Melting point, °C	98	650	660	1410	44	119	-101
Boiling point, °C	892	1107	2450	2680	280	445	-35

Many important engineering metals and alloys belong to the three transition series. Here, in addition to one or two electrons in the outermost *s* orbital, there are partially filled *d* orbitals. When bonding occurs, there is overlap of *s* orbitals as well as some overlap of *d* orbitals. As *d* orbitals have directional characteristics, the electronic structure of transition elements lends a partial covalent character to the bonding. This reflects in their properties, which fall between those of covalent crystals and typical metals. The thermal and electrical conductivities of transition metals are lower than those of typical metals such as copper and aluminium. The transition metals are also hard and not so ductile as copper or silver. The melting points of transition metals are higher than those of typical metals. Figure 4.9 shows the variation in bond energy, melting point, thermal expansion and density in the first transition series of elements.

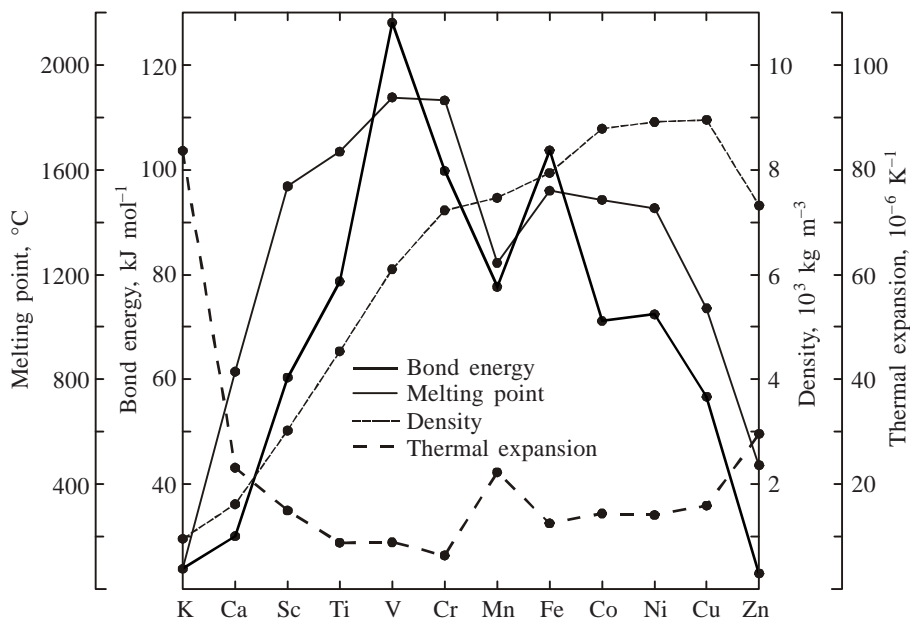


Fig. 4.9 The variation in the bond strength, melting point, thermal expansion and density of the metals of the first transition series.

SUMMARY

1. The atom consists of a nucleus and surrounding electrons. Each electron occupies a quantum state with a unique set of quantum numbers.
2. With increasing atomic number, the elements in the periodic table have increasing number of protons and electrons. The order of occupation of quantum states by electrons is determined by the Pauli exclusion principle, the Hund's rule, and the minimum energy criterion.
3. Ionization potential is the energy required to remove an electron from the outer orbital of an atom. Electron affinity is the energy released, when a free electron is added to the outer orbital. The electronegativity of an atom is a measure of its tendency to attract bonding electrons to itself.
4. The magnitude of the energy released, when two atoms come together from a large distance of separation to the equilibrium distance, is called the bond energy. It is related to the enthalpy of atomization of the solid. The centre to centre distance of the atoms at equilibrium is the bond length.
5. Primary bond energies are in the range 100–1000 kJ mol⁻¹. Secondary bond energies are in the range 1–50 kJ mol⁻¹.

6. Electrons can be transferred from an electropositive atom to an electronegative atom, producing ions of opposite sign and giving rise to the nondirectional ionic bond.
7. Sharing of electrons between neighbouring atoms results in a covalent bond, which is directional.
8. A metal is an array of positive ions which are held together in a cloud of free electrons. The metallic bond is nondirectional and generally weaker than ionic and covalent bonds.
9. Melting and boiling points of materials increase with increasing bond strength. Strong and directional bonds result in hard and brittle solids. Free electrons are responsible for the high thermal and electrical conductivities of metals.

PROBLEMS

- 4.1 Find the minimum uncertainty in determining the position of a particle, if the uncertainty in its momentum Δp is not to exceed 10^{-30} kg m s⁻¹.
Answer: 1.055×10^{-4} m.
- 4.2 Using the rest mass of the electron and the mass of a cricket ball, calculate the uncertainty in the velocity of each for $\Delta p = 10^{-30}$ kg m s⁻¹.
Answer: 1 m s⁻¹ and $\sim 10^{-29}$ m s⁻¹.
- 4.3 Calculate the limits within which the energy difference of an electronic transition should be, in order that the emitted or absorbed radiation is in the visible range (3900 to 7800 Å) of the electromagnetic spectrum.
Answer: 5.09×10^{-19} J (3.18 eV) to 2.55×10^{-19} J (1.59 eV).
- 4.4 Prepare a table similar to Table 4.2, showing the outer electron configuration of the second transition series of elements.
- 4.5 On the basis of values permissible for the four quantum numbers, derive the number of quantum states corresponding to the fourth principal shell ($n = 4$).
Answer: 32.
- 4.6 Give the electronic configuration of the fluorine atom and the F⁻ ion. Compare their sizes.
- 4.7 The heat of dissociation of the chlorine molecule is 121.3 kJ mol⁻¹. The Cl–Cl bond energy is 242.4 kJ mol⁻¹. Reconcile this difference.
- 4.8 Derive the units of the constant A' in Eq. (4.4). Compare it with the units of the dielectric constant in Chap. 17.

2. The order of increasing energy of orbitals in the sixth row of the Periodic Table is
A. $6s4f5d6p$ B. $6s6p4f5d$ C. $6s6p6d6f$ D. $6s5f5d6p$
3. If an electron transition occurs across the energy gap in GaAs of 1.43 eV, the radiation emitted or absorbed has a wavelength (Planck's constant = 6.626×10^{-34} J s; velocity of light = 2.998×10^8 m s⁻¹)
A. 86.7 Å B. 12400 Å C. 8670 Å D. 13890 Å
4. Nickel, which is to the left of Cu ($3d^{10}4s^1$) in the first transition series, has an outer electron configuration
A. $3d^94s^1$ B. $3d^84s^2$ C. $3d^{11}4s^1$ D. $3d^{10}4s^2$
5. If there are six electrons in the d orbital of a transition metal, the number of unpaired electrons are
A. 6 B. 5 C. 4 D. 0
6. The electron affinity of He in kJ mol⁻¹ is
A. 369 B. -704 C. 86 D. 0
7. Primary bonds have energy range in kJ mol⁻¹
A. 1000–5000 B. 100–1000 C. 10–100 D. 1–10
8. If the Fe–Fe bond length is 2.48 Å, the radius of the iron atom is
A. 2.48 Å B. 1.75 Å C. 1.43 Å D. 1.24 Å
9. The tetrahedral bond angle of (sp^3) bonds is
A. 90° B. 99° C. 104° D. 109.5°
10. If copper has bond energy of 56 kJ mol⁻¹ of bonds, the enthalpy of atomization of copper in the same units is about
A. 56 B. 112 C. 336 D. 672
11. Thermal expansion of materials arises from
A. strong bonds C. thermal vibrations
B. weak bonds D. asymmetry of potential energy curve
12. Hydrogen bonds are stronger than
A. van der Waals bonds C. ionic bonds
B. metallic bonds D. covalent bonds
13. If the radius of anion is r_a and of cation is r_c , the bond length is
A. $(r_c + r_a)$ B. $\sqrt{3}(r_c + r_a)$ C. $\sqrt{3}/2(r_c + r_a)$ D. $r_a - r_c$

Answers

- | | | | | |
|-------|-------|-------|------|-------|
| 1. C | 2. A | 3. C | 4. B | 5. C |
| 6. D | 7. B | 8. D | 9. D | 10. C |
| 11. D | 12. A | 13. A | | |

Suggestions for Further Reading

M.F.C. Ladd, *Structure and Bonding in Solid State Chemistry*, Ellis Horwood, Chichester, UK (1979).

L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press, Ithaca (1960).

H.H. Sisler, *Electronic Structure, Properties and the Periodic Law*, Reinhold Publishing Corporation, New York (1963).

As already noted, solids can be either crystalline or noncrystalline. In this chapter, we first look at a general comparison between the crystalline and the noncrystalline states. Following a bond-wise classification, we then describe the structure of elements, solid solutions and simple compounds. The structure of silica and the silicates is covered separately in a section. Towards the end, the structure and the general characteristics of polymers are described in some detail.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Density	kilogram per cubic metre	kg m^{-3}	g/cm^3 , lb/cu.in.
Specific gravity	—	—	—
Packing efficiency	—	—	—

Note: The density of water is 1000 kg m^{-3} (1 g/cm^3). The specific gravity of a material is the ratio of its density to the density of water. For example, the density of iron is 7800 kg m^{-3} . Its specific gravity is 7.8. The densities of typical materials are given in Appendices I and II.

Constants

Avogadro's number N	$= 6.023 \times 10^{23} \text{ mol}^{-1}$
Atomic mass unit (amu)* (also called <i>atomic weight</i> unit)	$= (10^3 N)^{-1} = 1.660 \times 10^{-27} \text{ kg}$
*For example, the mass of the carbon atom	$= 12 \text{ amu}$ $= 12 \times 1.660 \times 10^{-27} \text{ kg}$ $= 1.992 \times 10^{-26} \text{ kg}$

5.1 The Crystalline and the Noncrystalline States

The number and kind of nearest neighbours, that an atom or an ion has in a solid, is nearly the same for both the crystalline and the noncrystalline forms of the solid. However, the noncrystalline structure does not have the *long range periodicity* characteristic of the crystalline state. In a crystal, any number of integral lattice translations would take us from an atom located at a lattice point to another identical atom located at a different lattice point. This is not true for the noncrystalline state. As long range periodicity is the basis of diffraction effects, the noncrystalline solids do not give rise to sharp diffraction patterns like crystals.

Several factors promote the formation of noncrystalline structures. When primary bonds do not extend in all directions, one-dimensional chain molecules or two-dimensional sheet molecules are formed. Such units have to be aided by secondary bonding forces to form a three-dimensional crystal. Consider a structure consisting of long chain molecules. In the molten state, the chains persist and are like a bowl of wriggling earthworms, see Fig. 5.1a. The “wriggling” refers to the *translational motion* of the chains past one another. On cooling, if the secondary bonding forces are not strong enough to exert themselves, the chains cannot get straightened out of the tangle to become the orderly, parallel arrangement of a crystal as in Fig. 5.1b. The translational freedom is gradually lost during cooling, till the noncrystalline, glassy state is reached, where only the vibrational degree of freedom remains. The “earthworms” are frozen in place in the solid in some random configuration similar to that in Fig. 5.1a.

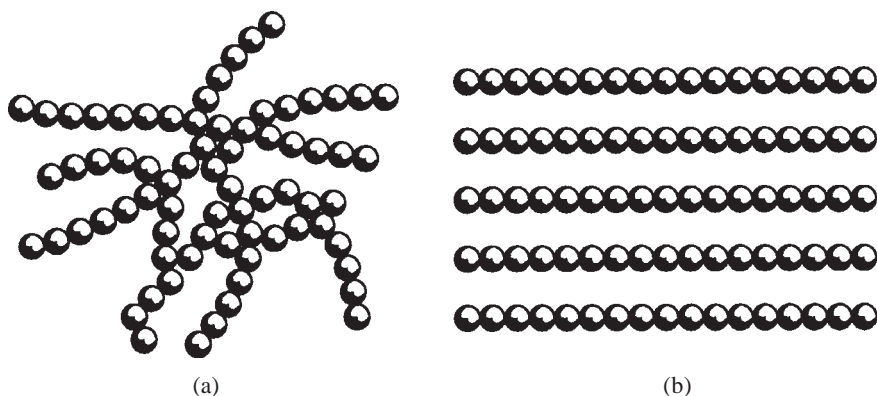


Fig. 5.1 (a) The tangled up configuration of long chain molecules, (b) the parallel array of the chain molecules characteristic of a crystal.

The free energy of the crystalline state is *always lower* than that of the noncrystalline state. Only when the free energy difference between the two states is large in magnitude, the tendency to crystallize will be strong. Some materials form a relatively open network structure of atoms, where there is little free

energy difference between an orderly array and a disorderly array of the units. In such cases, the tendency to crystallize will be weak.

A third factor that promotes the formation of noncrystalline structures is the rate of cooling from the liquid state. Kinetic barriers exist along the path of transition from the liquid to the crystalline state. Slow cooling rates allow enough time for crystallization, while fast cooling rates may prevent crystallization altogether. The rate of cooling which is considered as slow or fast will vary widely for different materials, depending on the magnitude of the kinetic barrier. For a metal, the cooling rate required to prevent crystallization may be as high as 10^6 K s^{-1} . For a silicate, cooling at the rate of a fraction of K per hour may be sufficient to prevent crystallization.

The crystal exhibits a *sharp melting point* in contrast to the noncrystalline material, which gradually softens over a range of temperature. As a result of the regularity of the arrangement, the atoms or molecules in a crystal are more closely packed. Hence, the crystalline form has a *higher density* than the noncrystalline form. The closer packing in the crystal tends to increase the average strength of the secondary bonds present.

INORGANIC SOLIDS

5.2 Covalent Solids

Starting from the *seventh column* of the periodic table, we note that the halogens are only one electron short to fill their outermost *p* orbitals and, therefore, each atom forms one covalent bond. This results in diatomic molecules of F_2 , Cl_2 , Br_2 and I_2 . Only weak secondary forces can hold these molecules together in a crystal. Fluorine and chlorine are in the gaseous state at ambient temperatures, bromine is a liquid, and iodine forms an orthorhombic crystal.

The *sixth column* elements form two covalent bonds, by the sharing of the two half-filled *p* orbitals of each atom. A number of these elements form *long zigzag chains* which are held together by secondary bonds in a solid. The zigzagging reflects the angular relationship between the *p* orbitals taking part in bonding. It is difficult to produce a stable three-dimensional structure by holding together long, one-dimensional chains with the aid of weak bonds. Hence, these elements are frequently in the noncrystalline form.

The two bonding electrons can also produce ring molecules such as S_8 , small molecules like H_2O and H_2S , or doubly bonded diatomic molecules such as O_2 . These molecules are bonded together by secondary forces when they form a crystal. For example, H_2O forms the hexagonal ice crystal with hydrogen bonds between the water molecules.

P, As, Sb and Bi of the *fifth column* with three half-filled *p* orbitals form three bonds. Their structure consists of *puckered sheets* in which each atom has three nearest neighbours. The angular relationships between the *p* orbitals are more or less preserved. The sheets are held together by van der Waals forces.

The intersheet bonding, even though stronger than the interchain bonding, is still weak, so that these materials are also found in the noncrystalline form. The covalent bond length within a sheet is smaller than the van der Waals bond length across two neighbouring sheets. This difference, however, decreases as we go down the column from phosphorus to bismuth, with the changing character of the primary bond. Antimony and bismuth form in the molten state nearly close packed structures characteristic of typical molten metals with as many as 10 or 11 nearest neighbours on an average. However, they crystallize with three covalently bonded neighbours to an atom. As a result of this, they *expand* on solidification, filling minute cavities and reproducing the mould details accurately, a property that is exploited in type casting. The three bonding electrons of the fifth column elements also produce small molecules such as NH_3 and triply bonded N_2 .

When each carbon atom forms three covalent bonds due to (sp^2) hybridization, *sheets of graphite* are produced. Unlike the puckered sheets of the fifth column elements, the graphite sheets are planar, as the three (sp^2) bonds are co-planar with an interbond angle of 120° . The sheets are held together in a crystal by van der Waals bonds as depicted in Fig. 5.2. The fourth bonding

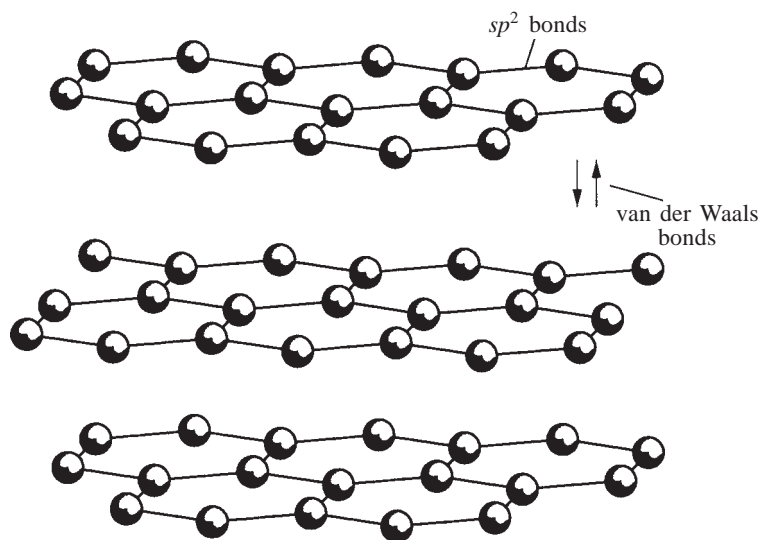


Fig. 5.2 Sheets of graphite are held together by secondary bonds in the crystal.

electron of carbon is delocalized and resonates between the three (sp^2) bonds. Its mobility accounts for a 100-fold increase in electrical and thermal conductivity in a direction parallel to the sheets, as compared to the perpendicular direction. The weak intersheet bonding explains the *softness* of graphite in sharp contrast to diamond, the other crystalline form of carbon, which is the hardest known mineral. Consequently, graphite is used as a lubricant. It is also the 'lead' in

pencils, where the softness is controlled by varying the proportion of clay to graphite, giving different grades of pencils, from 4B to 4H. Its high sublimation temperature and appreciable electrical conductivity are utilized in resistance heating applications. Also, if the sheets are aligned in a fibre such that the sheets are parallel to the fibre axis, the mechanical strength and the elastic modulus can be increased by orders of magnitude.

In the other crystal form of carbon, *diamond*, each atom forms four bonds, as a result of (sp^3) hybridization. These four bonds produce a three-dimensional network of primary bonds. Diamond exists in two crystal forms: the cubic and the hexagonal. In both the forms, the directions of the bonds from any atom in the network are given by the lines joining the centre to the corners of a regular tetrahedron, with an interbond angle of 109.5° . We will discuss only the better known *diamond cubic* (DC) structure. The DC unit cell is shown in Fig. 5.3a. A plan view of the positions of atoms in the unit cell is shown in Fig. 5.3b. The numbers indicated at atom positions represent the height of those positions from the base of the cube, the height of the unit cell being taken as unity. Two numbers at the same position denote two atoms, one above the other.

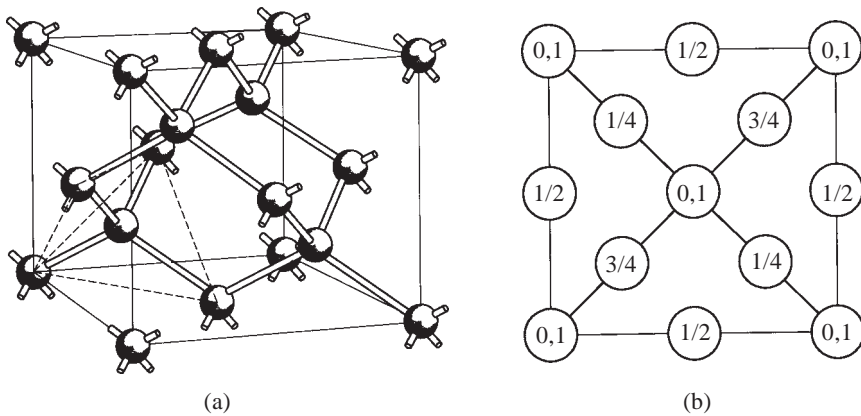


Fig. 5.3 (a) The diamond cubic (DC) unit cell, (b) plan view of atom positions in the unit cell. [The numbers indicate the height from the base of the cell, the total height of the cell being taken as unity.]

The space lattice of the DC crystal is FCC, with two atoms per lattice point. The basis has one atom at a lattice point, say, at one corner of the unit cell and the other atom at a point quarter way along the body diagonal, which is not a lattice point. The distance of separation between the two atoms of the basis, which are nearest neighbours, is $a\sqrt{3}/4$, where a is the lattice parameter. This distance (1.54 \AA) defines the diameter of the carbon atom in the DC crystal. The number of nearest neighbours of a carbon atom, known as the *coordination number*, is four. This low coordination, dictated by the covalent bonding, results in a relatively inefficient packing of the carbon atoms in the crystal.

Example 5.1 Calculate (i) the packing efficiency, and (ii) the density of diamond.

Solution

(i) Effective number of atoms in the DC unit cell

$$\begin{aligned}
 &= \frac{1}{8} \times 8 \text{ (corner atoms)} \\
 &\quad + \frac{1}{2} \times 6 \text{ (face centred atoms)} \\
 &\quad + 1 \times 4 \text{ (atoms fully within the unit cell)} \\
 &= 8.
 \end{aligned}$$

$$\text{Volume of each atom} = (4/3)\pi r^3 = (4/3)\pi (a\sqrt{3}/8)^3,$$

where r is the radius of the atom and a is the lattice parameter.

$$\begin{aligned}
 \text{Packing efficiency} &= \frac{\text{Volume of atoms in the unit cell}}{\text{Volume of unit cell}} \\
 &= \frac{8 \times (4/3)\pi (a\sqrt{3}/8)^3}{a^3} \\
 &= 0.34
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii) Density} &= \frac{\text{Mass of atoms in unit cell}}{\text{Volume of unit cell}} \\
 &= \frac{12 \times 1.660 \times 10^{-27} \times 8}{(3.57 \times 10^{-10})^3} \\
 &= 3500 \text{ kg m}^{-3} \\
 &= 3.5 \text{ g cm}^{-3}
 \end{aligned}$$

Recently, several new structures of carbon, known by the general name of *fullerenes*, have been discovered. One of them is the giant C_{60} molecule. Like graphite, it is made up of (sp^2) bonds, which are somewhat distorted to make the three-dimensional molecule. The structure of C_{60} molecule is illustrated in Fig. 5.4. The starting three-dimensional framework is the icosahedron, which has twenty equal and equilateral triangular faces as illustrated in Fig. 5.4a. By snipping off each of the twelve vertices of this figure, we obtain the (football-like) truncated icosahedron, see Fig. 5.4b. At each junction in Fig. 5.4b, a carbon atom is located, which is bonded to three neighbours. As compared to the planar graphite sheet, the three bonds here do not lie on a plane. In another related structure known as *carbon nanotubes*, the graphite sheets are rolled into a number of concentric hollow tubes, with the average diameter ranging from 20

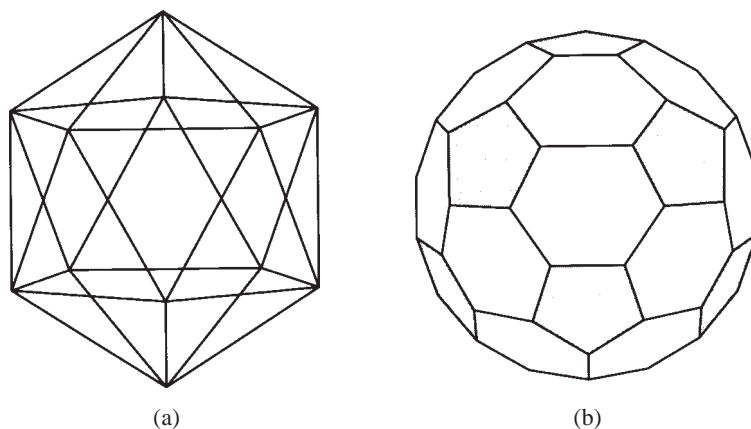


Fig. 5.4 Structure of C_{60} molecule: (a) the icosahedron; (b) the truncated icosahedron.

to 60 Å. The tubes can be up to 1 mm in length. As very small fibres of high strength, they have the potential to be a very effective reinforcing material for a softer matrix e.g. a polymer.

The other elements of the fourth column Si, Ge, and gray tin also have the diamond cubic structure of carbon. The lattice parameter of the DC type crystal increases with increasing atomic number as tabulated as follows:

Element	C (dia)	Si	Ge	Gray tin
Atomic number	6	14	32	50
Lattice parameter, Å	3.57	5.43	5.65	6.46

A large number of compounds with equal atomic fractions of two elements crystallize in forms closely related to the cubic and hexagonal forms of diamond. They are made up of two elements of IV column or one element each of III and V columns or II and VI columns or I and VII columns of the periodic table. Equal atomic fractions of these combinations give *on an average four electrons per atom* needed for the tetrahedral covalent bonding. Some examples are:

IV–IV Compound: SiC

III–V Compounds: AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, InSb

II–IV Compounds: ZnO, ZnS, CdS, CdSe, CdTe

I–VII Compounds: CuCl, AgI

For example, in cubic ZnS, the sulphur atoms are at the body corners and the face centres of the unit cell. The zinc atoms are at the $(1/4, 1/4, 3/4, 3/4)$ positions within the unit cell.

Si, Ge and the compounds listed above form the vast majority of semiconductor crystals. They are used in a number of solid state devices such as diodes, transistors, radiation detectors, photoelectric devices, solar batteries, thermistors and lasers.

Diamond, being very hard, is used in wire drawing dies and as an abrasive in polishing and grinding operations. Silicon carbide is cubic and is used as an abrasive and as heating element in furnaces.

5.3 Metals and Alloys

As metallic bonds are *nondirectional*, each metal atom in a crystal tends to surround itself with as many neighbours as possible, for minimizing the potential energy. As a first approximation, we can take the metal atoms to be hard, incompressible spheres and look at the *geometry of close packing* of such equal-sized spheres. Close packing is used here to mean the closest possible packing. Close packing along a row is obtained by arranging spheres in contact with one another along a row, as shown in Fig. 5.5a. Another row of close packed spheres can be placed against the first row, such that each sphere of the second row fits in the space between two adjacent spheres of the first row touching both of them, as illustrated in Fig. 5.5b. Repetition of this step produces a close packed plane, see Fig. 5.5c.

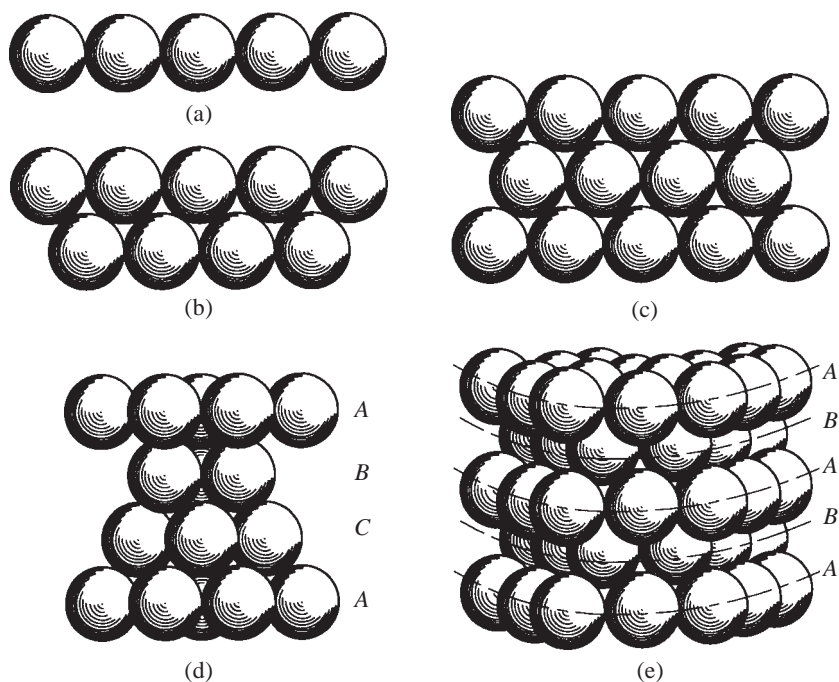


Fig. 5.5 Close packing of spheres (a) along a row, (b) on two adjacent rows, (c) on a plane, (d) three-dimensional ...ABCABC... stacking, and (e) ...ABABAB... stacking.

A three-dimensional block of close packed planes can be built in the following ways. Consider the configuration of a central sphere surrounded by six spheres in a close packed plane. The spheres are all labelled *A* and are shown slightly separated from one another for the sake of clarity in Fig. 5.6a. There is

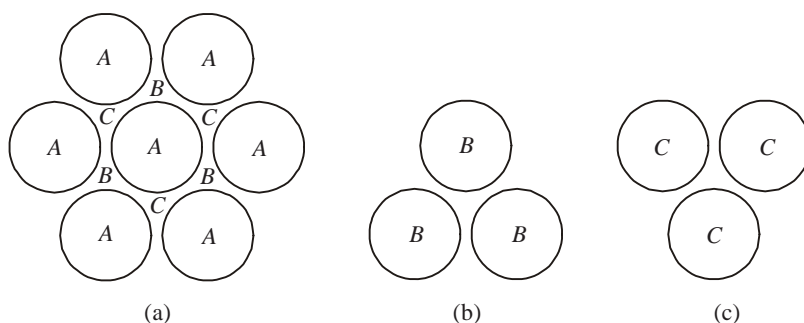


Fig. 5.6 (a) A central sphere surrounded by six spheres on a close packed plane forms two sets of alternate valleys *BBB* and *CCC*; (b) three spheres to fit in the set of alternate valleys *BBB*; (c) three spheres to fit in the set of alternate valleys *CCC*.

a triangular valley between any three neighbouring spheres of the plane. There are six such valleys between the central sphere and the six neighbours. A set of three alternate valleys out of the six can be chosen to fit the spheres of a second close packed plane. One set is marked *BBB* and the other *CCC* in Fig. 5.6a. Let us choose the set *BBB* to position the three second-plane spheres *BBB* (Fig. 5.6b) on top of the first. The other spheres of the second plane (not shown) would then automatically fit into appropriate valleys of the first plane. Let a third plane be now stacked such that any three of its spheres *CCC* (Fig. 5.6c) are vertically over the second set of valleys *CCC* of the first plane. These three planes of the stacking are laterally shifted with respect to one another, that is, the projections of the centres of the spheres on the base do not overlap. We can call these planes *A*, *B* and *C*. Note that no other lateral position such as *D* is possible. By repeating this sequence of stacking, we obtain the face centred cubic (FCC) stacking, see Fig. 5.5d,

... *ABCABCABCAB*...

as this gives rise to the FCC monoatomic crystalline arrangement.

Figure 5.7 shows the orientation of the close packed planes in the FCC unit cell. The neighbouring atoms are shown separated to improve clarity. The *ABCA* sequence is indicated. This figure should be compared with Fig. 3.3c, where the atoms are drawn to their correct sizes touching along the face diagonal. When the close packed planes are tilted to the horizontal position, the vertical direction will coincide with the body diagonal *XY* in Fig. 5.7.

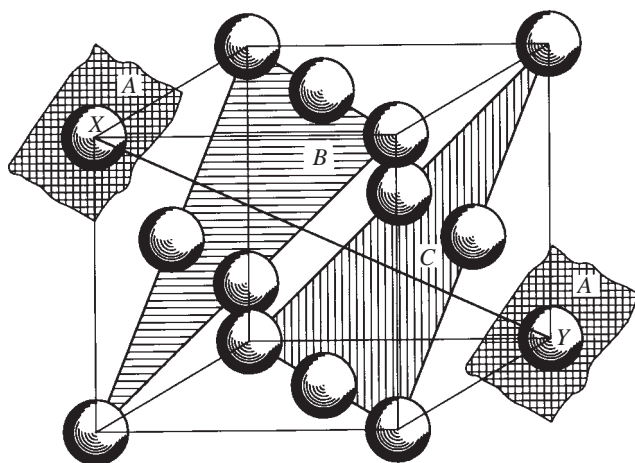


Fig. 5.7 Orientation of the close packed planes in the FCC unit cell.

Example 5.2 Determine the Miller indices of the close packed planes of an FCC crystal.

Solution From Fig. 5.7, it is seen that the close packed planes belong to the $\{111\}$ family. There are four pairs of these planes. The two planes comprising a pair are parallel. However, the pairs are nonparallel to one another.

- (i) (111) , $(\bar{1}\bar{1}\bar{1})$
- (ii) $(\bar{1}11)$, $(1\bar{1}\bar{1})$
- (iii) $(1\bar{1}1)$, $(\bar{1}1\bar{1})$
- (iv) $(11\bar{1})$, $(\bar{1}\bar{1}1)$

The FCC stacking can be visualized as parallel planes of anyone of the above four sets. The direction perpendicular to the planes will then coincide with one of the four nonparallel body diagonals of the cube.

In the FCC stacking, after choosing one set of valleys BBB for positioning the second plane, we used the other set CCC for the third plane. Instead, we could have positioned the spheres of third plane such that their projections will coincide with those of the first plane AAA . Repetition of this sequence gives rise to the *hexagonal close packed (HCP) stacking* (Fig. 5.5e),

$$\dots ABABABA \dots$$

as this sequence results in the HCP crystal.

The *unit cell* of the HCP crystal is shown in Fig. 5.8. For clarity, the close packed planes are shown separated. They are parallel to the hexagonal base of the unit cell. Hence, they are called the basal planes. The spheres on the top hexagon (A plane) are vertically above the spheres of the bottom hexagon (also A plane). The middle plane is a B plane with its spheres fitting in the triangular

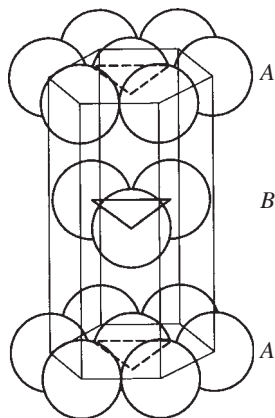


Fig. 5.8 The unit cell of the HCP crystal. For clarity, the close packed planes are shown separated.

valleys of the A planes. The A plane atoms are at lattice points, while the B plane atoms are not, refer to Table 3.1 for lattice point distribution in the hexagonal lattice. The effective number of lattice points is 3, whereas the effective number of atoms in the unit cell is 6, three of which are from the B plane. The *basis* therefore consists of two atoms per lattice point.

There are other ways of stacking close packed planes of spheres, the only restriction being that no two adjacent close packed planes can have the same symbol. A stacking such as ...ABBA... is not permissible. The *packing efficiency* of all close packed stackings is 0.74, that is, 26% of the space in the close packing of equal sized spheres is ‘empty’. The coordination number for all close packings is 12, that is, each sphere has 12 nearest neighbours, 6 in the same plane and 3 each in the two adjacent planes above and below.

Some crystals have arrangements of atoms which do not correspond to close packing. The most important of these is the arrangement in which each sphere has 8 nearest neighbours. This gives rise to the body centred cubic crystal, the unit cell of which is shown in Fig. 3.3b. The effective number of atoms or lattice points in the unit cell is 2. The atoms touch each other along the body diagonal. It can easily be shown that the packing efficiency of this arrangement is 0.68, only 6% lower than that for close packing.

The packing efficiency depends on the coordination number as shown in Table 5.1.

TABLE 5.1
Coordination Number and Packing Efficiency

Crystal	Coordination number	Packing efficiency
Diamond cubic (DC)	4	0.34
Simple cubic (SC)	6	0.52
Body centred cubic (BCC)	8	0.68
Face centred cubic (FCC)	12	0.74

Packing efficiencies higher than 0.74 are possible only with spheres of different sizes.

Example 5.3 Calculate the c/a ratio for an ideally close packed HCP crystal.

Solution The ...ABA... type of stacking represents the HCP structure depicted in Fig. 5.9. Joining the centres of the three neighbouring atoms of the

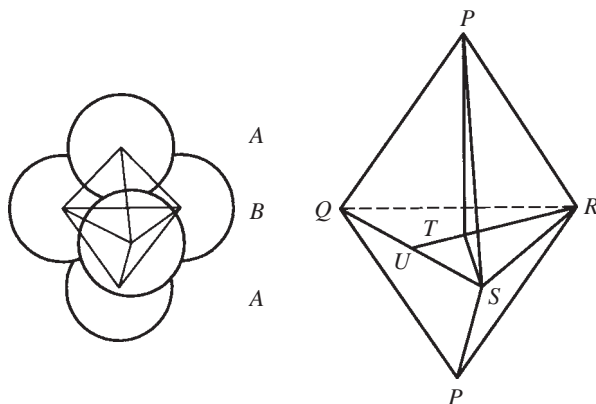


Fig. 5.9 In the ...ABA... packing, which is HCP, the centres of three atoms Q , R and S on plane B are joined to the centres of P atoms in plane A above and below.

middle plane to the centres of the atoms of the top and the bottom planes results in two tetrahedra with a common base. The top and the bottom atoms are centred at two lattice points, one above the other on the two hexagonal basal planes of the unit cell. So, the distance between them is the unit distance along the c -axis. The distance between any two adjacent atoms of a plane is unit distance along the a -axis. Unit of c is equal to twice the normal from the apex of a tetrahedron to its base. Unit of a is equal to the side of the tetrahedron.

In Fig. 5.9,

$$c/a = \frac{2PT}{RS}$$

$$RU = \sqrt{RS^2 - SU^2} = \sqrt{a^2 - a^2/4} = \sqrt{3} a/2$$

$$RT = \frac{2}{3} RU = a/\sqrt{3}$$

$$PT = \sqrt{PR^2 - RT^2} = \sqrt{a^2 - a^2/3} = a\sqrt{2}/\sqrt{3}$$

$$c/a = 2(a\sqrt{2}/\sqrt{3})/a = 1.633.$$

Any three-dimensional array of spheres gives rise to *void space* between the spheres. These voids are called *interstitial voids*. There are two main types of interstitial voids in close packed structures:

- tetrahedral voids
- octahedral voids.

A *tetrahedral void* is formed between three spheres on a close packed plane and a fourth sphere on an adjacent plane fitting in the cavity space between the three spheres—either on top (upright tetrahedral void, Fig. 5.10a) or at the bottom (inverted tetrahedral void, Fig. 5.10b). The name ‘tetrahedral void’ comes from

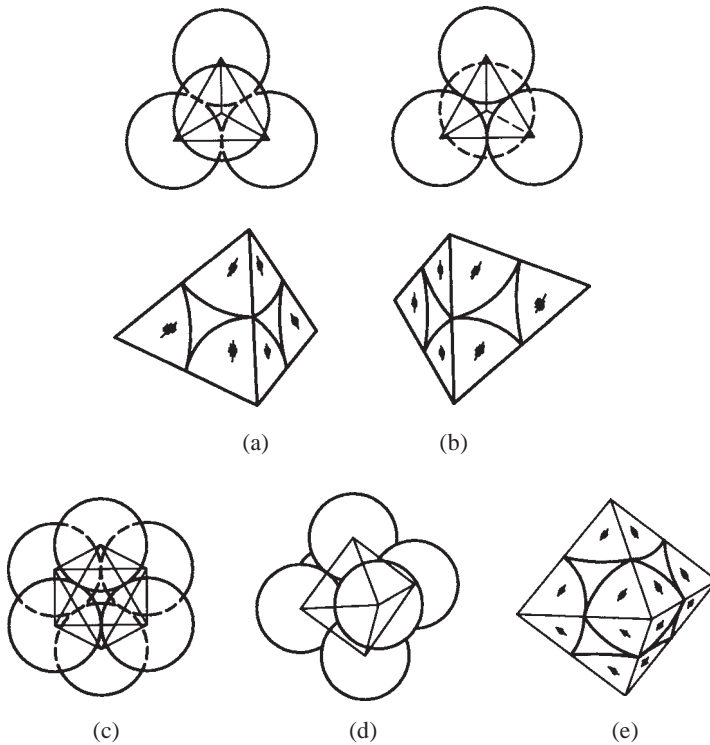


Fig. 5.10 Geometry of tetrahedral and octahedral voids in close packed structures.

the regular tetrahedron obtained by joining the centres of the four spheres. There are *two tetrahedral voids for every sphere* in the three-dimensional array. For example, in the FCC unit cell, the centres of the tetrahedral voids lie quarter-way and three-quarter-way along the four nonparallel body diagonals of the cube. There are thus eight tetrahedral voids in the unit cell. As the effective number of atoms in the unit cell is 4, this works out to a ratio 2:1 for the tetrahedral voids to the atoms.

Example 5.4 Find the size of the largest sphere that can fit into the tetrahedral void of a close packed structure without distorting it.

Solution Let r be the radius of the close packed spheres that form the tetrahedral void. Referring to Fig. 5.9, the distance from the corner of the tetrahedron to the centre is $3/4PT$. This distance is equal to the sum of the radii of the close packed sphere and the smaller sphere that fits in the tetrahedral void. So, the size of the sphere that fits the tetrahedral void is $3/4PT - r = \sqrt{3}r/\sqrt{2} - r = 0.225r$.

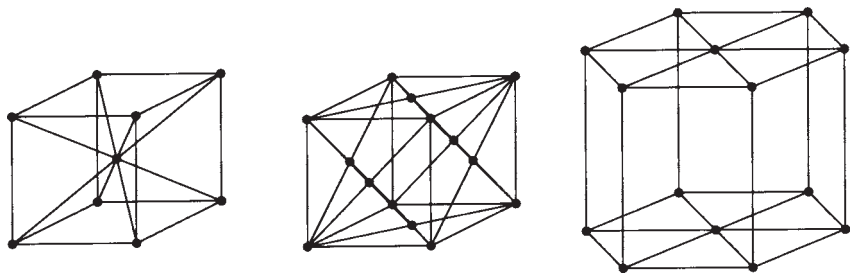
An *octahedral void* is formed with three spheres on a close packed plane and three more spheres on an adjacent close packed plane, as shown in Fig. 5.10c. Note that the three spheres of the adjacent plane are positioned such that the centres of the three spheres are directly over the three triangular valleys surrounding the central valley of the first plane, with no sphere over the central valley. Figure 5.10d is a tilted view of the octahedral arrangement, showing the square base with one sphere each on top and at the bottom. The name comes from the regular octahedron (a polyhedron with six corners and eight faces, the faces being equal and equilateral triangles) formed by joining the centres of the six spheres. There is one octahedral void per sphere in the three-dimensional array. For example, in the FCC unit cell, the centres of the octahedral voids fall at the body centre and at the middle of the 12 cube edges. The effective number of octahedral voids per unit cell is then $1 + 12 \times 1/4 = 4$. This is equal to the effective number of atoms in the unit cell. It can be shown that the largest sphere that can fit an octahedral void is $0.414r$, where r is the radius of the spheres of the close packed array.

The crystal structures of some common metals are given below. More details of the crystal structures and the lattice parameters of metallic elements can be obtained from the table on the back inside cover of the book.

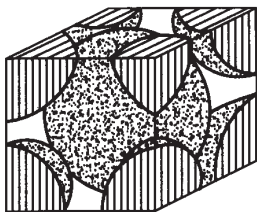
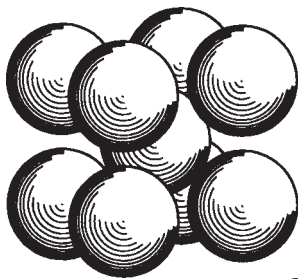
FCC	Ag	Al	Au	Cu	Fe (910 to 1410°C)			Ni	Pb	Pt		
BCC	Cr	Fe (below 910°C)			K	Li	Mo	Na	Nb	Ta	V	W
HCP	Be	Cd	Mg	Ti	Zn							

A number of metals have more than one crystal form. Iron, for example, is BCC at room temperature and changes over to the FCC form at 910°C. At 1410°C, iron again changes over to the BCC form. In general, at higher temperatures, the BCC crystal structure is to be expected as it allows larger vibrational amplitudes for atoms, thereby increasing the (thermal) entropy and lowering the free energy of the crystal. This is believed to be the reason for a number of alkali metals adopting the BCC form at room temperature, which is close to their melting range of 20–200°C. The partial covalent character of the transition metals is a possible reason for a number of them having the BCC structure at low temperatures.

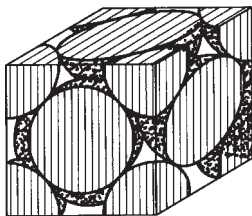
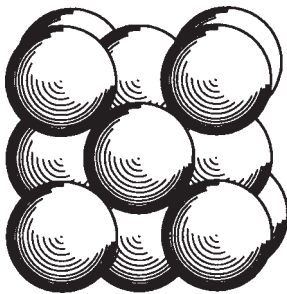
As a review, the space lattices, the unit cells and the sharing of atoms by the unit cells are illustrated below for BCC, FCC and HCP crystals:



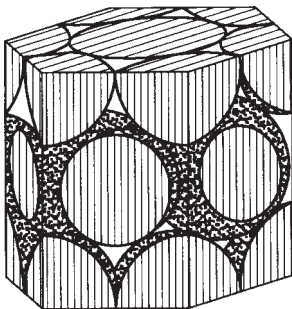
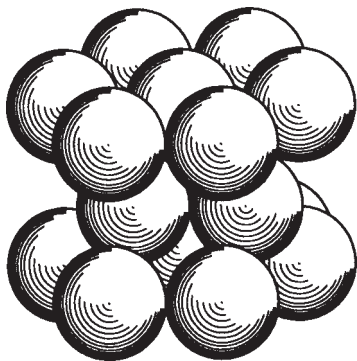
The space lattices



BCC unit cell



FCC unit cell



HCP unit cell

A number of metals dissolve in each other forming solid solutions. Solid solutions are analogous to liquid solutions. The mixing of the elements in the solid is on the atomic scale. When a solute atom is much smaller than the solvent atom, it may dissolve interstitially occupying a void space in the parent structure. For example, carbon is an interstitial solute in FCC iron and occupies the octahedral voids in the FCC structure. When the solute and the solvent atoms are of comparable sizes, the solute substitutes for the solvent atom on a regular atomic site. For example, a 70% Cu-30% Zn alloy (alpha brass) has an FCC structure, with copper and zinc atoms occupying randomly the atomic sites of the FCC crystal. Hume Rothery has framed empirical rules that govern the formation of substitutional solid solutions. Extensive solid solubility by substitution occurs when

- (i) the solute and the solvent atoms do not differ by more than 15% in diameter;
- (ii) the electronegativity difference between the elements is small; and
- (iii) the valency and the crystal structure of the elements are the same.

Ag–Au, Cu–Ni and Ge–Si systems satisfy the Hume Rothery conditions very well as shown in Table 5.2. These systems therefore form complete solid solutions, that is, the two elements mix in each other in all proportions. Starting from pure silver, for example, the silver atoms can be continuously replaced by gold atoms in the FCC structure till pure gold is obtained.

TABLE 5.2
Parameters Relevant to the Hume Rothery Rules

System		Crystal structure	Radius of atoms, Å	Valency	Electro-negativity
Ag–Au	Ag	FCC	1.44	1	1.9
	Au	FCC	1.44	1	2.4
Cu–Ni	Cu	FCC	1.28	1	1.9
	Ni	FCC	1.25	2	1.8
Ge–Si	Ge	DC	1.22	4	1.8
	Si	DC	1.18	4	1.8

This obviously is not possible in the case of Cu and Zn which have FCC and HCP structures, respectively. They, therefore, form solid solutions up to a certain extent only. At room temperature, up to 35% of zinc can dissolve in the FCC crystal of copper, as the Hume Rothery conditions are partially satisfied. However, only about 1% of copper dissolves in the HCP structure of zinc. The reason for this difference in behaviour seems to lie in the fact that an excess of bonding electrons, for example, when zinc is dissolved in copper, is more easily accommodated than a deficiency of bonding electrons which is seen when copper is dissolved in zinc.

Substitutional solid solutions have usually a random arrangement of the constituent atoms on the atomic sites, especially at elevated temperatures. This is

so, as configurational entropy makes a greater contribution in lowering the free energy with increasing temperature, recall that $G = H - TS$. This random arrangement of the constituent atoms in a solid solution may change over to an ordered arrangement on cooling to lower temperatures if ordering lowers the enthalpy of the crystal sufficiently. For example, the solid solution of copper and zinc mixed in equal atomic proportions forms a disordered BCC structure at temperatures above 470°C. The atomic sites in the BCC crystal are fixed, but the probability of finding a given atomic site occupied by a copper atom or a zinc atom is 0.5, that is, in the same proportion as the concentration of the constituent atom. Below 470°C, the alloy becomes ordered, with all the copper atoms occupying the cube corners and the zinc atoms occupying the body centres (or vice versa). Such a structure is identical to the one shown in Fig. 3.4c. This ordering occurs, as there is some preference for Cu–Zn bonds which have a larger bond energy than Cu–Cu or Zn–Zn bonds. Note that in the ordered state, all the zinc atoms have copper nearest neighbours and all the copper atoms have zinc neighbours, that is, all the bonds are of the Cu–Zn type.

The above example of the 50 Cu–50 Zn alloy is an intermediate structure that forms in a system of limited solid solubility. Its crystal structure (BCC) is different from that of either copper (FCC) or zinc (HCP). If an intermediate structure exists only at a fixed composition, it is called an *intermetallic compound*. For example, iron carbide (Fe_3C), a common constituent of steels, is an intermetallic compound. It has a complex crystal structure referred to as an orthorhombic lattice and is hard and brittle.

Intermediate structures, which have appreciable difference in electronegativity of the constituent atoms, obey the normal *rules of valency*. Examples, where magnesium is one of the constituent atoms, are Mg_2Si , Mg_3P_2 , MgS , Mg_2Ge , and MgSe .

Some intermediate compounds, which do not obey the normal rules of valency, are called *electron compounds*. Hume Rothery has shown that they occur at certain definite values of free electron to atom ratio in the alloy such as 3 : 2, 21 : 13 and 7 : 4. Typical examples are CuZn (3 : 2), Cu_5Zn_8 (21 : 13) and CuZn_3 (7 : 4).

5.4 Ionic Solids

In Chapter 4, we noted that the size of an atom increases on adding extra electrons and shrinks on removing electrons. Consequently, *the cation is usually smaller than the anion*. Exceptions to this exist in a few cases such as RbF , with Rb^+ larger than F^- . Ionic bonds are *nondirectional*. Therefore, each ion tends to surround itself with as many ions of opposite sign as possible to reduce the potential energy. This tendency promotes the formation of close packed structures. Unlike the metallic structures, here the differences in size of the two (or more) ions forming the crystal should also be taken into account in the geometry of close packing.

Consider first the *local packing geometry* of one type of cation and one type of anion. The cation is assumed to be the smaller ion. The number of anions surrounding a central cation is called the coordination number or *ligancy*. The

ligancy is a function of the ion sizes and can be worked out from space filling geometry when the following conditions corresponding to a stable configuration are satisfied *simultaneously*:

- (i) An anion and a cation assumed to be hard spheres always touch each other.
- (ii) Anions generally will not touch, but may be close enough to be in contact with one another in a limiting situation.
- (iii) As many anions as possible surround a central cation for the maximum reduction in electrostatic energy.

When the cation is very small compared to the anion, it is easily seen that only *two* anions can be neighbours to the cation in order to satisfy all the above three conditions. Consider next the configuration shown in Fig. 5.11a. Here, the three surrounding anions are touching one another and also the central cation.

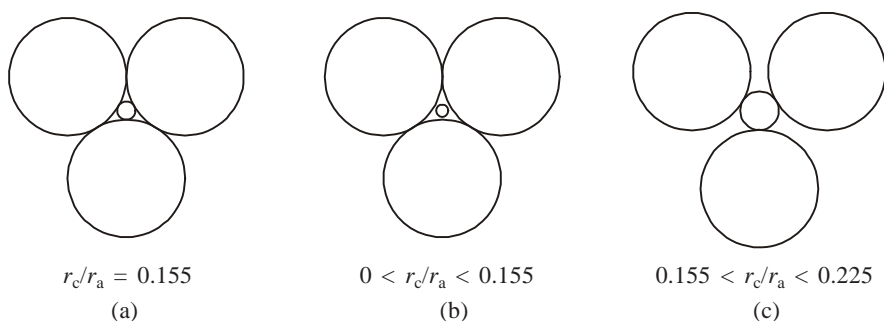


Fig. 5.11 Triangular coordination of anions around a central cation: (a) the critical configuration, (b) the unstable configuration; and (c) stable (but not critical) configuration.

The ratio of the cation to anion radius r_c/r_a for this configuration is 0.155, which can be worked out from the simple geometry (Example 5.5). The *triangular arrangement* in Fig. 5.11a is one of the limiting situations. The radius ratio is said to be a critical value because for values of r_c/r_a smaller than 0.155, the central cation will rattle in the hole and not touch all the three anions at the same time, as illustrated in Fig. 5.11b. This violates condition (i) above and leads to instability. When the radius ratio is less than 0.155, the only way to satisfy all three conditions is to reduce the number of anions to 2. For values of r_c/r_a slightly greater than 0.155, all the anions touch the central cation but do not touch one another, as shown in Fig. 5.11c. All three conditions of stability are still satisfied. This situation will prevail till the radius ratio increases to 0.225, the next higher critical value corresponding to a *tetrahedral (four) coordination*. At $r_c/r_a = 0.225$, the four surrounding anions touch one another and also the central cation. This configuration is the same as that obtained by fitting the largest possible sphere in the tetrahedral void of a close packed structure, see Example 5.4.

A ligancy of five does not satisfy all the three conditions for stable configuration because it is always possible to have six anions as an alternative to any arrangement that contains five anions, without a change in the radius ratio. The critical condition for *octahedral (six) coordination* occurs at $r_c/r_a = 0.414$, which is the same as the size of the octahedral void in a close packed structure. Ligancies of 7, 9, 10 and 11 are again not permissible. The radius ratio ranges in which different values of ligancy are obtained are summarized in Table 5.3. At the end of the table, the limiting case of $r_c/r_a = 1$ is identified with configurations of close packing of equal sized spheres.

TABLE 5.3
Ligancy as a Function of Radius Ratio

Ligancy	Range of radius ratio	Configuration
2	0.0 – 0.155	linear
3	0.155 – 0.225	triangular
4	0.225 – 0.414	tetrahedral
6	0.414 – 0.732	octahedral
8	0.732 – 1.0	cubic
12	1.0	“FCC” or “HCP”

Example 5.5 Find the critical radius ratio for triangular coordination.

Solution The critical condition for triangular coordination is shown in Fig. 5.11a. The three anions touch one another as well as the central cation. From the simple geometry, we can write

$$\begin{aligned}
 r_c + r_a &= \frac{2}{3} (2r_a) \sin 60^\circ \\
 &= 1.155r_a \\
 r_c/r_a &= 0.155.
 \end{aligned}$$

The ligancy rules outlined above are obeyed in a number of cases. For example, in the NaCl crystal, the radius ratio $r_{\text{Na}^+}/r_{\text{Cl}^-} = 0.54$, which lies between 0.414 and 0.732. As listed in Table 5.3, the predicted ligancy is six. The octahedral geometry of six chlorine ions surrounding a central cation is experimentally observed. In MgO, $r_{\text{Mg}^{2+}}/r_{\text{O}^{2-}} = 0.59$ and again the octahedral coordination is observed. In CsCl, $r_{\text{Cs}^+}/r_{\text{Cl}^-} = 0.91$. This value lies in the range of 0.732–1.0. The predicted coordination of eight anions surrounding a cation is observed. In a marginal case such as CaF_2 , where $r_{\text{Ca}^{2+}}/r_{\text{F}^-} = 0.73$, it is difficult to predict whether a six-fold or an eight-fold coordination will occur. It so turns out that the eight coordination is observed in this case, that is, every calcium cation is surrounded by eight fluorine anions.

The Si–O bond in silica as well as in silicates is about 50% ionic and 50% covalent. Here, the central silicon cation is surrounded by four oxygen anions located at the corners of a regular tetrahedron. This arrangement satisfies *both* the ligancy rules (as $r_c/r_a = 0.29$, the tetrahedral coordination is predicted from Table 5.3), as well as the orientation relationships of (sp^3) bonds.

The stability criteria listed above for predicting the ligancy may not always be valid. If directional characteristics of bonding persist to any significant degree, the considerations based on the radius ratio alone will not lead to the correct prediction of ligancy. In the above-discussed example of Si–O coordination, the radius ratio criterion and the bond angle requirements happen to coincide. In ZnS, where the bond is more covalent than ionic, the ligancy predicted from $r_{Zn^{2+}}/r_{S^{2-}} = 0.48$ is octahedral. Yet the four-fold coordination characteristic of (sp^3) bonding is what is observed.

In the formation of ionic crystals, the ligancy rules described above determine the local packing around a cation. The long-range arrangement of ions in the crystal is dependent on the following factors:

- (i) In the crystal, the *overall electrical neutrality* should be maintained, whatever be the net charge on a local group of a cation and surrounding anions. For example, in NaCl, where a cation is surrounded by six anions, the net charge on $(NaCl)_6$ is five. Evidently, this has to be neutralised in the long range arrangement.
- (ii) The ionic bond being nondirectional, the ions are packed as closely as possible in the crystal, consistent with the local coordination.
- (iii) If small cations with a large net charge are present in the crystal, the cation–cation repulsion will be high. Then it may be necessary to have a long range arrangement that maximizes the cation–cation distance, even if close packing is not possible. Such a situation usually arises when the charge on the cation increases to three or four.

When the cation charge is not more than two or at best three and when the radius ratio is in the range 0.414–0.732, the crystal structure can be described as a FCC or HCP packing of anions with the cations occupying all or part of the octahedral voids in the structure. The fraction of octahedral voids that are filled depends on the number of cations to anions in the chemical formula. Thus, for the *rock salt* (NaCl) structure, adopted by hundreds of binary ionic compounds, $r_{Na^+}/r_{Cl^-} = 0.54$, and the anion packing is FCC with all octahedral voids filled with sodium cations. Recall that there is one octahedral void per sphere in a close packed array. A unit cell of NaCl crystal is shown in Fig. 5.12, with the larger chlorine ions occupying the face centred cubic positions and the sodium ions in the octahedral voids. The octahedral positions are at the body centre and at the midpoints of the cube edges. Note that, unlike in the monoatomic FCC crystal (refer Fig. 3.3c), the chlorine ions do not touch one another along the face diagonal. This is so because the radius ratio of 0.54 is greater than the size of the octahedral void in a close packed structure, which is 0.414. The FCC close packing is ‘opened up’ here to the extent necessary to accommodate the sodium cations in the octahedral voids.

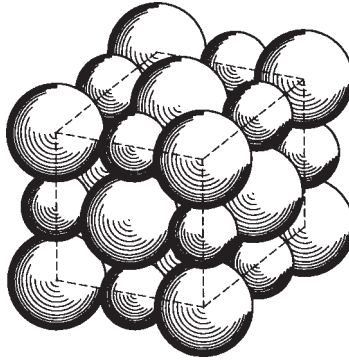


Fig. 5.12 Unit cell of NaCl with chlorine ions at the FCC positions and the sodium ions in the octahedral voids.

As the FCC positions and the octahedral void centres are interchangeable like the body centre and the body corners in the BCC cell, the NaCl structure can also be described as two interpenetrating monoatomic FCC cells, one corresponding to the anions and the other to the cations. The basis of the NaCl structure is one sodium ion plus one chlorine ion. The sum of their radii, $r_{\text{Na}^+} + r_{\text{Cl}^-} = a/2$, where a is the lattice parameter.

Example 5.6 Calculate the density of MgO from the following data.

Structure: FCC anion packing, cations in the octahedral voids.

Radius of Mg^{2+} ion = 0.78 \AA

Radius of O^{2-} ion = 1.32 \AA

Solution MgO has the same structure as NaCl. So,

$$\begin{aligned} \text{the lattice parameter } a &= 2(r_{\text{Mg}^{2+}} + r_{\text{O}^{2-}}) \\ &= 2(0.78 + 1.32) = 4.20 \text{ \AA} \end{aligned}$$

The effective number of oxygen anions at FCC positions in the unit cell = $8 \times 1/8$ (corner ions) + $6 \times 1/2$ (face centred ions) = 4.

The effective number of magnesium cations in the octahedral voids = 1 (body centre) + $12 \times 1/4$ (midpoints of cube edges) = 4.

$$\begin{aligned} \text{Density} &= \frac{\text{Mass of atoms in unit cell}}{\text{Volume of unit cell}} \\ &= \frac{(16 + 24.3) \times 1.660 \times 10^{-27} \times 4}{4.20^3 \times 10^{-30}} \\ &= 3610 \text{ kg m}^{-3} \\ &= 3.61 \text{ g cm}^{-3} \end{aligned}$$

Several ionic crystals, which have the radius ratio in the range of octahedral coordination, are listed in Table 5.4. In alumina (Al_2O_3), the cation positions in neighbouring planes of octahedral voids are staggered such that the mutual repulsion of the trivalent cations is minimized. Due to multivalent ions, the bond strength in Al_2O_3 is high, producing a hard crystal with a high melting point.

TABLE 5.4
Structure of Some Ionic Crystals

Crystal	r_c/r_a	Anion packing	Fraction of octahedral voids with cations
NaCl	0.54	FCC	All
MgO	0.59	FCC	All
CdCl_2	0.57	FCC	Half
Al_2O_3	0.43	HCP	Two-thirds

The electrical insulating properties of Al_2O_3 are excellent. Al_2O_3 is used as a substrate for building integrated circuits and in spark plugs of automobiles. In the Al_2O_3 structure, it is possible to replace part or all of the Al^{3+} ions by other ions, provided the size difference between them is small. The structure thus produced is a *substitutional solid solution* already discussed in connection with alloys. Replacing a small fraction of the Al^{3+} ions by other ions such as Cr^{3+} and Fe^{3+} results in the gemstones, ruby and blue sapphire. Ruby (Cr^{3+} ions added to Al_2O_3) is a LASER (Light Amplification through Stimulated Emission of Radiation) which is a crystal device. Sapphire is very hard and is used in jewelled bearings and cutting tools.

When r_c/r_a is in the range 0.732–1, the eight-fold coordination is observed. CsCl with $r_{\text{Cs}^+}/r_{\text{Cl}^-} = 0.91$ is a typical example of this structure. Referring to Fig. 3.4c, the cesium ions are at the body centre and the chlorine ions are at the body corners. The space lattice is simple cubic, with a basis of one cesium ion plus one chlorine ion per lattice point. In CaF_2 also, the coordination around the cation is eight, with the difference that only one body centre for every two unit cells is occupied by Ca^{2+} .

As an example of an ionic crystal with more than two types of ions, consider the crystal structure of spinels. *Spinel*s are compounds with two different cations A^{2+} and B^{3+} and oxygen as the anion, with the general formula, AB_2O_4 . Here, the oxygen anions form the FCC packing. For every four oxygen anions, there are four octahedral sites and eight tetrahedral sites. Out of these twelve, only three are needed to fill the cations of the above formula. In the *normal spinel structure*, the A cations are in the tetrahedral voids and the B cations are in the octahedral voids. Alternatively, half of the B cations can occupy the tetrahedral voids while the remaining half of the B cations and all the A cations are randomly distributed in octahedral voids, resulting in the *inverse spinel structure*. In both normal and inverse spinels, only half of the octahedral sites and 1/8 of the tetrahedral sites are filled. When $\text{B}_2\text{O}_3 = \text{Fe}_2\text{O}_3$, we have a series of compounds called *ferrites*, where different A cations can be present in varying proportions. Ferrites have the inverse spinel structure. They are important soft magnetic materials (refer Chap. 16).

5.5 The Structure of Silica and the Silicates

The silicate tetrahedron, shown in Fig. 5.13, is a basic repeating unit in the structure of silica and the silicates. A tetravalent silicon ion is surrounded by

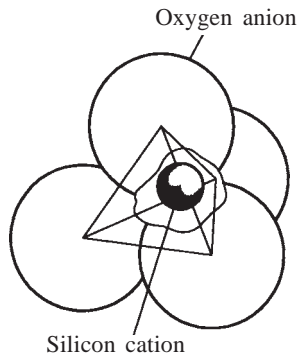


Fig. 5.13 The silicate tetrahedron has a silicon cation at the centre of the tetrahedron and four oxygen anions at the four corners.

four oxygen anions, the formula of the unit being $(\text{SiO}_4)^{4-}$. The four excess negative charges on the unit should be neutralised by the formation of primary bonds with other units (or other cations). The cation–cation repulsion is high here. Therefore, silica (SiO_2) cannot form a close packed layer of oxygen anions with cations in the tetrahedral voids. Such a structure brings neighbouring silicon cations so close to one another that it results in an appreciable increase in the potential energy due to the repulsion between them.

The structure of silica is shown in Fig. 5.14. It forms a three-dimensional network of tetrahedra, each one of which shares all its four corners with other

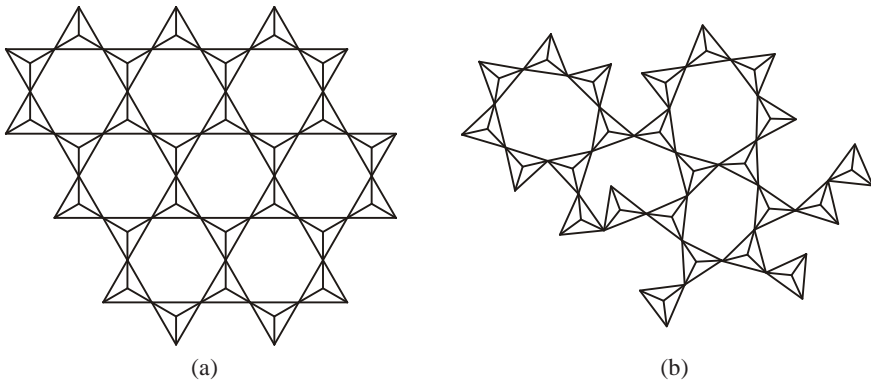


Fig. 5.14 Structure of silica in (a) crystalline and (b) noncrystalline forms.

tetrahedra, that is, the oxygen anions at the corners are common to two tetrahedra. The effective number of silicon cation per tetrahedral unit is 1, and the effective number of oxygen per unit is $4 \times 1/2 = 2$, as each corner is shared

by two tetrahedra. This arrangement maintains the electrical neutrality of the network as a whole. In the crystalline form such as *quartz*, the tetrahedra are arranged in a periodically repeating pattern, see Fig. 5.14a. In the noncrystalline form, *silica glass*, the tetrahedra are randomly bonded to other tetrahedra, as shown in Fig. 5.14b. Note that there is no difference in the silicon-oxygen coordination between the crystalline and the noncrystalline forms.

Quartz is used in optical components. It is also *piezoelectric* (piezo = press). A mechanical stress applied to the crystal displaces the ions in the crystal and induces electric polarization. Similarly, an electric field will cause the crystal to be elastically strained. Quartz is used in watches and clocks and for accurate frequency control in electronic circuits. Fused silica glass is used in applications requiring low thermal expansion. It is highly viscous even in the molten state because of the Si–O bond between the tetrahedra.

In the three-dimensional network of silica, other oxides can be dissolved to yield a number of both crystalline and noncrystalline silicates. *Soda lime glass* is a noncrystalline silicate with Na_2O and CaO added to silica. The alkali cations break up the network of the silicate tetrahedra, as shown in Fig. 5.15. For each Na_2O introduced, one Si–O bridge is disrupted and the extra oxygen atom from Na_2O splits up one common corner into separate corners. The two sodium ions

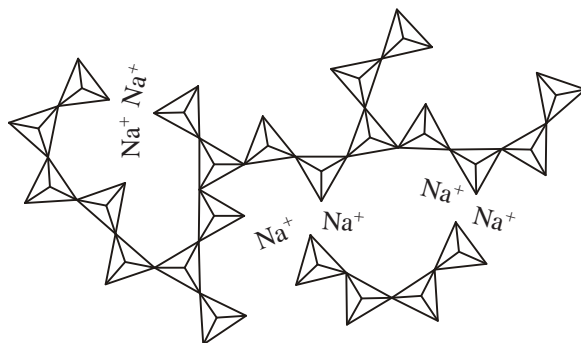

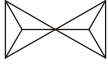

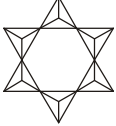
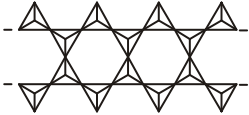
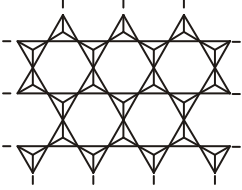


Fig. 5.15 Addition of Na_2O to silica introduces weaker bonds in the network.

stay close to the disrupted corner due to the electrostatic attraction. The network at the corner is bonded through the Na–O bonds. The Na–O bond being weaker than the Si–O bond, the viscosity of the glass is drastically reduced as a result of the alkali addition. Along similar lines, we can explain the much lower softening temperature of *pyrex* (80% SiO_2 , 14% B_2O_3 , 4% Na_2O) as compared to *fused silica* (99.8% SiO_2).

When the tetrahedra share all corners, the only cation in the structure is the silicon at the centre of the tetrahedron. By introducing other cations, a number of different structures can be produced. These can be described by reference to the number of corners the tetrahedra share amongst themselves without an intervening link provided by a different cation. According to the corners shared, the structures of many minerals can be classified as island, chain, sheet and three-dimensional network of tetrahedra, as shown in Table 5.5.

TABLE 5.5
Silicate Structural Units

Number of oxygen ions shared	Structural unit	Structural formula	Charge balance	Examples of minerals
0		Island (SiO_4) ⁴⁻ (ortho)	Si + 4 O - 8 Net - 4	Olivine (Mg,Fe) ₂ SiO ₄
1		Island (Si_2O_7) ⁶⁻ (pyro)	Si + 8 O - 14 Net - 6	Hemimorphite Zn ₄ Si ₂ O ₇ (OH) ₂ · H ₂ O
2		Single (SiO_3) ²⁻ chain	Si + 4 O - 6 Net - 2	Enstatite MgSiO ₃
		Ring (SiO_3) ²⁻	Si + 4 O - 6 Net - 2	Beryl Be ₃ Al ₂ (SiO ₃) ₆
2 $\frac{1}{2}$		Double (Si_4O_{11}) ⁶⁻ chain	Si + 16 O - 22 Net - 6	Tremolite (Asbestos) Ca ₂ Mg ₅ (OH) ₂ · (Si ₄ O ₁₁) ₂
3		Sheet (Si_2O_5) ²⁻	Si + 8 O - 10 Net - 2	Muscovite (Mica) KAl ₂ (OH) ₂ · (Si ₃ Al)O ₁₀
4	Three-dimensional network	(SiO ₂) ⁰	Si + 4 O - 4 Net 0	Quartz SiO ₂

When the tetrahedra do not share any corner directly, we have the *island structure*, that is, at all the four corners, each tetrahedron is joined to other tetrahedra only through other cations in the mineral. A *single chain structure* is

produced when each tetrahedron is joined at two corners to two other tetrahedra directly (that is, by having a common oxygen), with the remaining two corners joined through other cations. This results in some interesting *cleavage* properties of the minerals. A cleavage direction refers to the direction along which the bonds are broken during fracture. The fibrous quality of *asbestos* (which has a double chain structure) is attributable to the cleavage along certain crystallographic directions which go through the weaker bonds of the structure. Similarly, *mica* with tetrahedra arranged in sheet form breaks parallel to the sheets. The bond between sheets of *clay* is van der Waals, which explains the softness of clay. The excess charge on one side of the sheet compared to the other side provides ideal sites for absorption of the polar water molecules. So, clay has the characteristic plasticity when mixed with water. *Talc* is soft, like clay, due to van der Waals bonding between sheets but does not absorb water like clay.

Cement and *concrete* are common building construction materials used in huge quantities. Portland cement consists of a number of silicate minerals with the approximate composition given in Table 5.6. Cement mixed with water sets as a function of time due to several hydration reactions. Water binds the sheet-like silicate molecules together, thereby hardening the cement. Concrete is a sized aggregate of rocks embedded in a cement matrix, which binds the rock pieces together.

TABLE 5.6
Approximate Composition of Portland Cement

Constituent	Symbol	Percentage
Tricalcium silicate	C ₃ S	45
Dicalcium silicate	C ₂ S	30
Tricalcium aluminate	C ₃ A	10
Tetracalcium aluminoferrite	C ₄ AF	8
Other bonding agents	—	7

POLYMERS

Most organic polymers are based on the covalent bonds formed by carbon. The electrons are bonded strongly by the localized sharing, characteristic of covalent bonding so that polymers are good thermal and electrical insulators. The relative inertness of polymeric materials (they do not corrode in the sense that metals do) can be attributed to covalent bonding.

5.6 Classification of Polymers

The word ‘mer’ in Greek means a unit, and so *monomer* stands for a single unit and *polymer* for many units joined together. A polymer usually has thousands of

monomers bonded by a chemical reaction. For this reason, it is often called a *macromolecule*. If joining occurs essentially along one direction, we produce *long chain polymers*. These long chains are held together by secondary bonds in a polymeric solid. If joining occurs in other directions as well, we may produce a network of primary bonds yielding a three-dimensional giant molecule.

Polymers are broadly classified as

- (i) thermoplasts, and
- (ii) thermosets,

depending on their behaviour with a change in temperature. *Thermoplasts* have the property of increasing plasticity, that is, increasing ability to deform plastically with increasing temperature. Thermoplasts are long chain molecules held together by secondary bonds. As the thermal energy increases, the secondary bonds between the chains are broken more readily, rendering the plastics easily mouldable and ultimately melting it. This high degree of plasticity is one of the technologically attractive properties of plastics. At the same time, this easy melting at temperatures of the order of a few hundred degrees Celsius renders the plastics unsuitable for high temperature applications or for high strength applications at room temperature.

Thermosets have a three-dimensional network of primary bonds as polymerization proceeds in all directions. They are, therefore, relatively hard and rigid at room temperature and do not soften on heating. In fact, they become harder due to the completion of any left-over polymerization reaction. However, on heating, before reaching the melting temperature, they degrade, that is, they decompose owing to their reaction with atmospheric oxygen. *Bakelite* used for electrical insulation is a thermoset. The smell of bakelite in a short circuited electrical appliance is a sign of degradation.

Long chain polymers can be subdivided into three types:

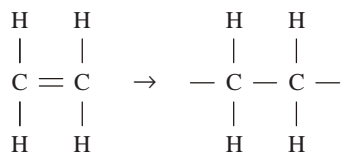
- Plastics
- Fibres
- Elastomers.

The long chain molecules are randomly oriented in a *plastic*. The plastic has therefore no directional property, unless it has been mechanically worked in such a way as to align the chains. In *fibres*, all the chains are more or less aligned in the long direction of the fibre. This alignment gives the fibres unique directional properties. Their strength and elastic modulus in a direction parallel to the fibre length are an order of magnitude larger as compared to the nonaligned structure. The third type of polymers are the *elastomers* which are long chain molecules exhibiting the unique rubbery behaviour.

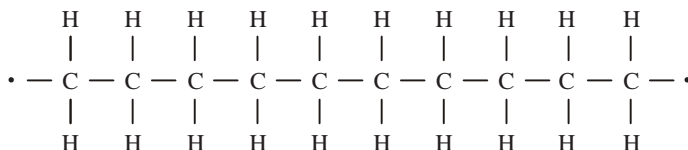
5.7 Structure of Long Chain Polymers

The formation of a long chain polymer starting from a monomer is easily understood by considering the *ethylene molecule* C_2H_4 , which has a double bond

necessary for chain formation. During polymerization, the double bond is opened up into two single bonds:



The monomers are bonded together end-to-end in a polymerization reaction:



In the bond representation above, it must be borne in mind that each carbon atom is *tetrahedrally bonded* to four neighbours with an (sp^3) bond angle of 109.5° , as shown in Fig. 5.16.

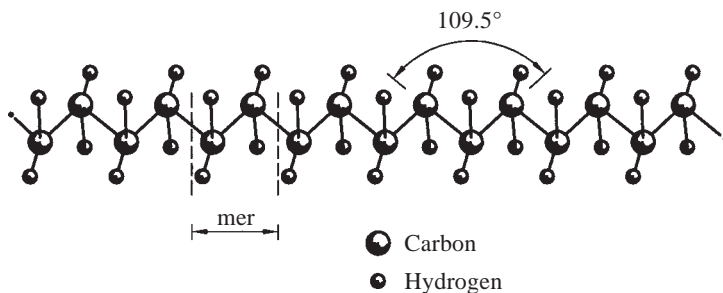


Fig. 5.16 The structure of a polyethylene chain.

The *degree of polymerization* (D.P.) defines the number of repeating monomers in a chain. The molecular weight of the chain molecule is equal to D.P. multiplied by the molecular weight of the monomer. Molecular weights of chains typically range from 10 000 to 1 000 000.

Table 5.7 lists the structure of typical ethylene-based long chain polymers. The structure is described by giving the formula of the four side groups (or radicals) denoted by R_1 , R_2 , R_3 and R_4 . A number of familiar plastics such as polyethylene, polypropylene, PVC and PTFE are ethylene based, as shown in Table 5.7.

TABLE 5.7
Ethylene Based Long Chain Polymers

Name	Monomer structure				Uses
	R ₁	R ₂	R ₃	R ₄	
Polyethylene (polythene)	H	H	H	H	Sheets, tubes and containers
Polyvinyl chloride (PVC)	H	H	H	Cl	Electrical insulation, gramophone records
Polypropylene	H	H	H	CH ₃	Ropes and filaments, vacuum flasks, flash light casings
Polymethylmethacrylate (Plexiglass)	H	H	CH ₃	COOCH ₃	Transparent windows and fixtures
Polystyrene	H	H	H	C ₆ H ₅	As styrofoam, sound proofing in refrigerators and buildings
Polytetrafluoroethylene (PTFE) (Teflon)	F	F	F	F	Coating for frying pans, razors and bearings, chemical ware, human body implants
Polyacrylonitrile (Orlon)	H	H	H	CN	Acrylic fibre used as wool substitute in clothings

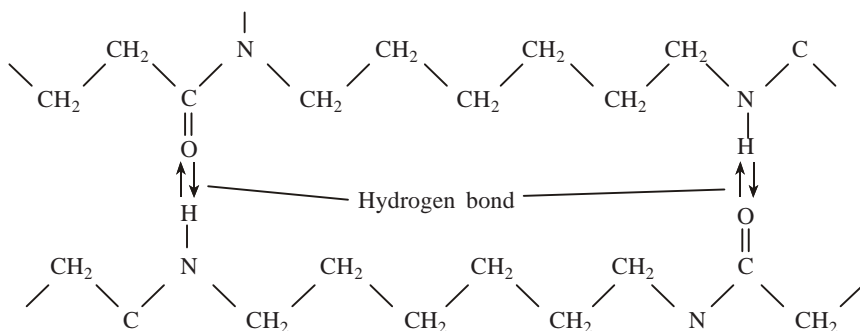
Many fibres are based on monomers which are not ethylene derivatives. The

polyamide fibres are characterized by $\left[\begin{array}{c} \text{O} \\ || \\ -\text{C}-\text{N}- \\ | \\ \text{H} \end{array} \right]$ linkage.

Nylon (6,6), which belongs to this category, has the following monomer formula:



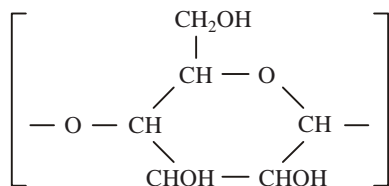
Here, six carbon atoms are present in the backbone of the chain between two neighbouring (NH) groups. The oxygen in the side group of the linkage provides for a bond with a hydrogen of an adjacent chain, as illustrated now.



Hydrogen bonds, being stronger than van der Waals bonds, account for the greater resistance of nylons to softening with increase in temperature. The

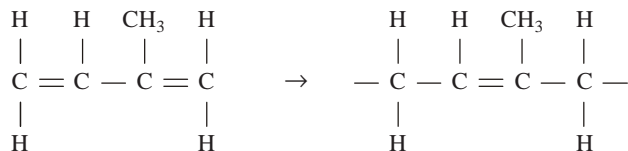
polyester fibres are characterized by the $\left[\text{— O — } \overset{\text{O}}{\parallel} \text{C —} \right]$ linkage. The fibre *terylene* (or dacron) comes under this category. Note that, in contrast to the polyamide linkage, there is an oxygen here in place of carbon in the backbone of the chain. This oxygen provides flexibility, so that a polyester softens easily with increasing temperature, in spite of the hydrogen bonding between chains.

The main constituent of *wood* is the cellulose chain, with the following monomer formula:



Alignment of the long chains gives wood its highly *directional* properties. The elastic modulus and the tensile strength are some 10–20 times more in the longitudinal direction as compared to the transverse direction. *Rayon*, or regenerated cellulose, is natural cellulose rendered highly crystalline. *Cotton* is a related natural plant fibre.

The structure of elastomers can be understood with reference to *natural rubber*. The monomer here is the *isoprene molecule*. The side groups in the monomer are both on the same side of the molecule. This arrangement results in a natural tendency for the molecule to bend and promotes rubbery behaviour. One of the two double bonds is used in chain formation:



The remaining double bond is necessary for producing an elastomer. Natural rubber, as tapped from a tropical tree, is a viscous liquid, in which the long polyisoprene molecules are able to flow past one another at room temperature. An elastomer is produced by heating raw rubber with sulphur. Sulphur forms covalent bonds with the carbon, by using the remaining double bond in the monomer. This reaction, known as *vulcanization*, produces additional links between chains called *cross links*. The degree of vulcanization determines the stiffness of rubber. With increasing cross links, the rubber becomes more rigid; for example, a cycle tube is less vulcanized than a rubber pocket comb. In the limit, when all the double bonds are used up by sulphur bridges, a three-dimensional network of primary bonds results. The material thus produced, *ebonite*, is hard and brittle. The elastic properties of rubber are discussed in Chap. 10.

5.8 Crystallinity of Long Chain Polymers

Long chain polymers are usually in the noncrystalline form or in the semicrystalline form. In the noncrystalline form, the long chains are randomly tangled with one another. In the semicrystalline form, parts of the polymer volume have the parallel chain arrangement while other parts are randomly oriented, as shown in Fig. 5.17. Sometimes it is possible to grow single crystals of a polymer. These crystals have a *folded chain structure*, where the same chain folds back and forth many times into a parallel arrangement. For example, the single crystal of polyethylene has an orthorhombic unit cell.

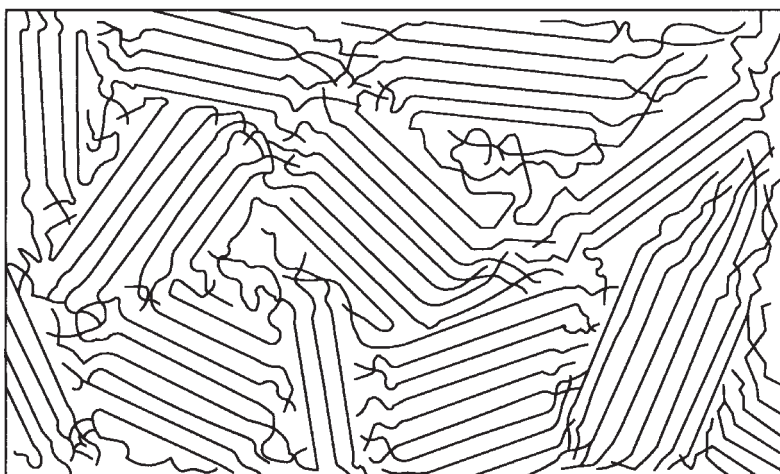


Fig. 5.17 Many polymers are semicrystalline, with random and parallel arrangements of chains in different regions.

In all polymers, the long chains can be aligned to some extent by mechanical working. Such an alignment promotes crystallinity. As the chains are more closely packed in the crystalline form, the density increases with increasing alignment of

the chains. Thus low density polyethylene (sp. gr. 0.92) is less than 50% crystalline, as compared to high density polyethylene (sp. gr. 0.97) which is about 80% crystalline. Crystallinity is also promoted by the formation of a hydrogen bond between chains in place of the weaker van der Waals bond. For example, nylon and cellulose crystallize owing to the hydrogen bond.

There are several factors which promote noncrystallinity. Chains that are very *long* get tangled more easily and are therefore more difficult to crystallize. Another factor which impedes crystallization is the phenomenon of *branching*. A long chain at some point along its length can bifurcate into two branches. This introduces primary bond bridges between chains at random points and impedes the parallel arrangement of chains necessary for crystallization. Irradiation of a polymer can knock off a small side group such as hydrogen, where the end of another molecule can be bonded, producing a branched structure. Branching makes the polymer stiffer.

In many polymers, the four side groups are not the same. In polystyrene, for example, one of the side groups is the phenyl group, which is large and bulky, as compared to the other three groups which are hydrogen atoms. During polymerization, different arrangements along the chain length are possible. If the large side groups are randomly arranged on either side of the chain (*atactic arrangement*), crystallization is obviously difficult, as this arrangement does not allow neighbouring chains to come close to one another all along the length of the chains. On the other hand, an *isotactic arrangement*, where all the bulky side groups are on the same side, or the *syndiotactic arrangement*, where the bulky side groups alternate on either side of the chain in a regular fashion, makes it possible for the chains to be packed closely and uniformly, promoting crystallinity.

When two or more monomers are polymerized together, the monomers can get bonded in a random sequence along the length of the chain, producing what are known as *copolymers*. These are the polymeric analog of solid solutions discussed earlier. Clearly, copolymers promote noncrystallinity. If a polymer contains low molecular weight additives called *plasticizers*, chains are impeded in coming together by the presence of the extraneous molecules, and crystallization is difficult.

SUMMARY

1. In a noncrystalline solid, the nearest neighbours of an atom may be the same as in a crystal, but it lacks the long range periodic arrangement characteristic of a crystal.
2. A three-dimensional tetrahedral network of covalent bonds is found in elements of the fourth column of the periodic table and in compounds which, on an average, have four electrons per atom in the outer orbital. Many semiconductor crystals fall under this category.

3. More than half of the elements are metals with simple crystal structures such as FCC, HCP or BCC.
4. Many ionic crystals can be described as an FCC or HCP packing of anions, with the cations in the interstitial voids of the structure.
5. The structure of silica and the silicates can be described with reference to the basic unit, the silicate tetrahedron.
6. Polymers are broadly classified as thermoplasts and thermosets. The thermoplasts comprise plastics, fibres and elastomers.
7. The crystallinity of long chain polymers is influenced by a number of factors such as chain length, branching, presence of large random sidegroups, and plasticizers.

PROBLEMS

- 5.1 Draw a regular tetrahedron inside a cubic unit cell, such that its corners touch four corners of the cube. From this figure, find the tetrahedral angle between lines joining the centre to the four corners of the tetrahedron.
- 5.2 Referring to the diamond cubic unit cell of Fig. 5.3, explain why the second set of positions ($1/4, 1/4, 3/4, 3/4$) are vacant, even though there is enough space for four more carbon atoms to fit into these positions.
- 5.3 Give the coordinates of atom positions in cubic diamond and ZnS.
- 5.4 Calculate the effective number of atoms in the HCP unit cell. Describe the basis for generating the HCP structure starting from the hexagonal space lattice.
- 5.5 Calculate the packing efficiency of (i) close packed structures (FCC and HCP) (ii) monoatomic BCC and SC crystals.
- 5.6 Give the Miller indices of the family of close packed directions in SC, BCC, FCC and DC crystals, indicating the magnitude of the closest distance of approach between neighbouring atoms.
Answer: $\langle 100 \rangle$, $\langle 111 \rangle$, $\langle 110 \rangle$, none. a , $a\sqrt{3}/2$, $a/\sqrt{2}$ and $a\sqrt{3}/4$.
- 5.7 Calculate the void space in closest packing of n spheres of radius 1.000, n spheres of radius 0.414, and $2n$ spheres of radius 0.225.
Answer: 19%.
- 5.8 Find the diameter of the largest atom that would fit an interstitial void in FCC nickel without distortion.
Answer: 1.03 \AA .
- 5.9 Find the size of the largest sphere that will fit an interstitial void in a BCC crystal as a function of the atomic radius r . The void is located at $(0, 1/2, 1/4)$ and other equivalent positions.
Answer: $0.29r$.

- 5.10** Calculate the critical radius ratio for tetrahedral, octahedral and eight-fold coordination around a central cation in an ionic crystal.
- 5.11** Distinguish between atom sites and lattice points in a monoatomic FCC crystal and a NaCl crystal.
- 5.12** State the differences and common points between (i) CsCl, (ii) a monoatomic BCC crystal, and (iii) a monoatomic SC crystal.
- 5.13** What are the factors that determine the density of a crystal? Which of these factors is dominant in determining the decreasing order of density with increasing atomic number in the following elements?
Cu ($Z = 29$) 8960, Ge ($Z = 32$) 5320, and Se ($Z = 34$) 4790 kg m⁻³.
- 5.14** Aluminium has an FCC structure. Its density is 2700 kg m⁻³ (sp.gr. 2.7). Calculate the unit cell dimension and the atomic diameter.
Answer: 4.05 Å, 2.86 Å.
- 5.15** Iron changes from BCC to FCC form at 910°C. At this temperature, the atomic radii of the iron atoms in the two structures are 1.258 Å and 1.292 Å, respectively. What is the per cent volume change during this structural change?
Answer: -0.45%.
- 5.16** X-ray analysis of a Mn-Si alloy with 75 atomic per cent of Mn and 25 atomic per cent of Si showed that the unit cell is cubic and the lattice parameter $a = 2.86$ Å. The density of the alloy is 6850 kg m⁻³. Find the number of atoms in the unit cell.
Answer: 2.
- 5.17** Calculate the density of the CsCl crystal from the radii of the ions: Cs⁺ = 1.65 Å and Cl⁻ = 1.81 Å.
Answer: 4380 kg m⁻³.
- 5.18** Find the local packing arrangement in the ionic crystal CaO. Assuming the crystal structure to be cubic, calculate the density of the crystal. Given the radii: Ca²⁺ = 0.94 Å and O²⁻ = 1.32 Å.
Answer: Octahedral, 4030 kg m⁻³.
- 5.19** The bonds in diamond are predominantly covalent in character and so are the bonds along the chains of a long chain polymer. Why does the polymer melt at a much lower temperature?
- 5.20** The melting point of a polymer increases with increasing molecular weight of the chain. Explain why this is so.
- 5.21** Calculate the end-to-end distance of an uncoiled chain molecule of polyethylene that has 500 mers in it. The C-C bond length is 1.54 Å.
Answer: 1258 Å.
- 5.22** The degree of polymerization of a polystyrene chain is 10 000. Calculate the molecular weight of the chain.
Answer: 1 040 000.

- 5.23** By considering the bond energies involved, show that, during polymerization of ethylene monomers, there is a decrease in the enthalpy of the system.
- 5.24** What is the weight of sulphur that is required to fully cross link 68 kg of polyisoprene (ebonite) starting from natural rubber?
Answer: 32 kg.
- 5.25** Determine the shortest centre-to-centre distance between any two atoms among the four atoms located inside the diamond-cubic unit cell.
Answer: $a/\sqrt{2}$
- 5.26** In the DC unit cell of silicon ($a = 5.43 \text{ \AA}$),
 (a) derive the number of atoms per unit area of the $\{110\}$ planes, and
 (b) derive the number of atoms per metre along $\langle 111 \rangle$ directions.
Answer: $7.2 \times 10^{18} \text{ m}^{-2}$; $2.1 \times 10^9 \text{ m}^{-1}$
- 5.27** The Bragg angle for the first reflection from SiC is 17.2° . Determine the Bragg angle for the second reflection.
 [Hint: A weak reflection will be observed for $(h^2 + k^2 + l^2) = 4$, as there are two types of atoms in the DC unit cell.]
Answer: 20.0° .
- 5.28** In cubic ZnS, the radii of the Zn and S atoms are 0.83 and 1.74 \AA , respectively. Determine the lattice parameter.
Answer: 5.94 \AA .
- 5.29** What is the distance from an atom of an A plane to the nearest atom of a B plane in an ideal HCP crystal?
Answer: a (lattice parameter).
- 5.30** The density of cubic manganese is 7440 kg m^{-3} . The lattice parameter $a = 8.92 \text{ \AA}$. Find the number of atoms in the unit cell. List the possible space lattices.
Answer: 58, BCC, SC.
- 5.31** Find the radius of the largest sphere that will fit an interstitial void in magnesium (HCP). $a_{\text{Mg}} = 3.21 \text{ \AA}$. Take the c/a ratio of Mg to be the ideal value.
Answer: 0.66 \AA .
- 5.32** Find the diameter of the largest sphere that will fit the void at the centre of the cube edge of a BCC crystal.
Answer: $0.134a$ (a is lattice parameter).
- 5.33** Make neat sketches of a random substitutional solid solution and an interstitial solid solution, stating their probabilities of formation as a function of the atomic sizes involved.
- 5.34** An elemental crystal has a density of 8570 kg m^{-3} . The packing efficiency is 0.68. Determine the mass of one atom if the closest distance of approach between neighbouring atoms is 2.86 \AA .
Answer: 93.01 amu.

- 5.35** Calculate the density and packing efficiency of the NaCl crystal.
Answer: 2230 kg m^{-3} , 0.66.
- 5.36** List the space lattice, basis, number of ions of each type in the unit cell, and lattice parameter-ionic radii relationships for the cubic ionic crystals MgO and CsCl. The cation-to-anion ratios are 0.59 and 0.91, respectively.
- 5.37** Distinguish between thermoplastics and thermosets, using any three of their characteristics.
- 5.38** On the basis of their structure, describe how elastomers are different from the other long-chain polymers.
- 5.39** List the factors that promote noncrystallinity in long-chain polymers.

MULTIPLE CHOICE QUESTIONS

- In the DC unit cell of silicon ($a = 5.43 \text{ \AA}$), the number of atoms per metre along the body diagonal are
 A. $8.5 \times 10^9 \text{ m}^{-1}$ B. $2.1 \times 10^9 \text{ m}^{-1}$
 C. $2.6 \times 10^9 \text{ m}^{-1}$ D. $1.84 \times 10^9 \text{ m}^{-1}$
- The average distance between atoms along the body diagonal of the DC crystal is
 A. $a\sqrt{3}/4$ B. $a\sqrt{3}/8$ C. $a\sqrt{3}/2$ D. $a\sqrt{3}$
- If the lattice parameter of Si = 5.43 \AA and the mass of Si atom is $28.08 \times 1.66 \times 10^{-27} \text{ kg}$, the density of silicon in kg m^{-3} is
 A. 2330 B. 1115 C. 3445 D. 1673
- The close packed direction in Ge crystal is
 A. [100] B. [110] C. [111] D. none of these
- The lattice parameter of GaAs (radius of Ga = 1.22 \AA , As = 1.25 \AA) is
 A. 5.635 \AA B. 5.704 \AA C. 5.774 \AA D. 4.94 \AA
- In cubic ZnS (II-VI compound), if the radii of Zn and S atoms are 0.83 \AA and 1.74 \AA , the lattice parameter of cubic ZnS is
 A. 11.87 \AA B. 5.94 \AA C. 5.14 \AA D. 2.97 \AA
- The number of atoms along the body diagonal of the diamond cubic unit cell is
 A. 1 B. 2 C. 3 D. 4
- The close packed directions that lie on a (111) plane of an FCC crystal are
 A. [101], $[\bar{1}10]$, $[0\bar{1}\bar{1}]$ B. $[11\bar{1}]$, $[\bar{1}\bar{1}1]$
 C. $[\bar{1}00]$, $[0\bar{1}0]$, $[00\bar{1}]$ D. $[\bar{1}0\bar{1}]$, $[1\bar{1}0]$, $[011]$

9. An elemental crystal has a density of 8570 kg m^{-3} . The packing efficiency is 0.68. If the closest distance between neighbouring atoms is 2.86 \AA , the mass of one atom is ($1 \text{ amu} = 1.66 \times 10^{-27} \text{ kg}$)
 A. 186 amu B. 93 amu C. 46.5 amu D. 43 amu
10. Tick the close-packing arrangements in the following:
 A. ...ABABABA ... B. ...ABCABCABCA...
 C. ...ABABCBCABC... D. ...ACCBCABCABC...
11. The packing efficiency of a simple cubic crystal with an interstitial atom exactly fitting at the body centre is
 A. 0.52 B. 0.68 C. 0.73 D. 0.91
12. The basis of Zn crystal (HCP) contains (l.p. = lattice point)
 A. 1 Zn atom/l.p. B. 2 Zn atoms/l.p.
 C. 4 Zn atoms/l.p. D. none of these
13. Sb has a hexagonal unit cell with $a = 4.307 \text{ \AA}$ and $c = 11.273 \text{ \AA}$. If its density is 6697 kg m^{-3} and its atomic weight is 121.75 amu, the number of atoms per unit cell (hexagonal prism) is ($1 \text{ amu} = 1.66 \times 10^{-27} \text{ kg}$)
 A. 3 B. 6 C. 12 D. 18
14. The atomic fraction of tin in bronze (FCC) with a density of 7717 kg m^{-3} and a lattice parameter of 3.903 \AA is (at.wt. Cu = 63.54, Sn = 118.7; $1 \text{ amu} = 1.66 \times 10^{-27} \text{ kg}$)
 A. 0.01 B. 0.07 C. 0.10 D. 3.8
15. The diameter of the largest sphere that fits the void at the centre of a cube edge of a BCC crystal of lattice parameter a is
 A. $0.293a$ B. $0.414a$ C. $0.134a$ D. $0.336a$
16. Expressed as a function of atom radius r , the radius of the void at the midpoint of the edge of a BCC crystal is
 A. $0.36r$ B. $0.414r$ C. $0.15r$ D. $0.19r$
17. Assuming the ideal c/a ratio for HCP Ti, the radius of the largest sphere that will fit interstitially in Ti ($a = 2.95 \text{ \AA}$) is
 A. 0.53 \AA B. 0.61 \AA C. 0.66 \AA D. 1.22 \AA
18. An octahedron has
 A. 8 corners B. 8 faces C. 8 edges D. 12 edges
19. The number of tetrahedral voids in HCP unit cell (hexagonal prism) is
 A. 4 B. 6 C. 8 D. 12
20. In Hume-Rothery rules for extensive solid solubility, the atomic diameter of the solute and the solvent atoms should not differ by more than
 A. 50% B. 15% C. 2% D. 0%

21. The density of the CsCl crystal with ionic radii of $\text{Cs}^+ = 1.65 \text{ \AA}$ and $\text{Cl}^- = 1.81 \text{ \AA}$ is (at.wt. Cs = 132.9, Cl = 35.45; $1 \text{ amu} = 1.66 \times 10^{-27} \text{ kg}$)
A. 4383 kg m^{-3} B. 3373 kg m^{-3} C. 2386 kg m^{-3} D. 6746 kg m^{-3}
22. The fraction of octahedral voids filled by Al^{3+} ions in Al_2O_3 ($r_{\text{Al}^{3+}}/r_{\text{O}^{2-}} = 0.43$) is
A. 0.43 B. 0.287 C. 0.667 D. 1
23. If the lattice parameter of CsCl is 3.995 \AA , the bond length (between bonding species) is
A. 3.46 \AA B. 6.92 \AA C. 1.73 \AA D. 5.65 \AA
24. The packing efficiency of a NaCl crystal (radius of $\text{Na}^+ = 0.98 \text{ \AA}$, $\text{Cl}^- = 1.81 \text{ \AA}$) is
A. 0.52 B. 0.66 C. 0.68 D. 0.74
25. In ruby, Cr^{3+} ions substitute for
A. Al^{3+} ions B. Fe^{3+} ions C. Mg^{2+} ions D. O^{2-} ions
26. In CaF_2 with the eight-fold cubic coordination, the fraction of voids at the body centre occupied by Ca^{2+} ions is
A. 1 B. 0.5 C. 0.25 D. 0
27. The length of an uncoiled polyethylene chain with 500 mers is (C–C bond length = 1.54 \AA)
A. 500 \AA B. 770 \AA C. 1540 \AA D. 1258 \AA
28. Apart from three H atoms, the fourth side group in the monomer of polypropylene is
A. Cl B. CH_3 C. C_6H_5 D. COOCH_3
29. The side groups in polyvinyl chloride are
A. H, H, H, H B. H, H, H, Cl C. H, H, CH_3 , Cl D. H, Cl, Cl, Cl
30. The weight of sulphur in kg required to achieve 5% cross-linking of 68 kg of polyisoprene is
A. 1.6 kg B. 3.2 kg C. 16 kg D. 32 kg
31. The factors that promote noncrystallinity in polymers are
A. large random side groups B. branching
C. addition of plasticizers D. copolymerization
32. The degree of polymerization in polyethylene is 50,000. The average molecular weight in amu is
A. 700,000 B. 1,300,000 C. 1,400,000 D. none of these
33. In the PTFE (teflon) monomer, the four side groups are
A. F, F, F, F B. H, H, H, H C. H, H, H, Cl D. H, H, H, CH_3
34. The weight in amu of a polyethylene chain with 1000 mers is
A. 28,000 B. 280,000 C. 62,000 D. 42,000

35. The minimum number of double bonds required in the monomer for cross-linking is
 A. 0 B. 1 C. 2 D. 3
36. The bulkiest side group in the monomer is in
 A. teflon B. PVC C. PTFE D. polystyrene
37. The chemical formula of the isoprene molecule is
 A. $C_4H_5CH_3$ B. $C_3H_4CH_3Cl$ C. C_4H_3Cl D. $C_2H_3CH_3$

Answers

- | | | | | |
|----------------|-------|----------|-------|-------------|
| 1. B | 2. C | 3. A | 4. D | 5. B |
| 6. B | 7. B | 8. A, D | 9. B | 10. A, B, C |
| 11. C | 12. B | 13. D | 14. C | 15. C |
| 16. C | 17. B | 18. B, D | 19. D | 20. B |
| 21. A | 22. C | 23. A | 24. B | 25. A |
| 26. B | 27. D | 28. B | 29. B | 30. A |
| 31. A, B, C, D | 32. C | 33. A | 34. A | 35. C |
| 36. D | 37. A | | | |

Sources for Experimental Data

- C.A. Harper, *Modern Plastics Handbook*, McGraw-Hill, New York (2000).
- P. Villars and L.D. Calvert, *Pearson's Handbook of Crystallographic Data for Intermetallic Phases*, Vols. 1–3, American Society for Metals, Metals Park, Ohio (1985).

Suggestions for Further Reading

- R.C. Evans, *An Introduction to Crystal Chemistry*, Cambridge University Press, Cambridge (1964).
- A.F. Wells, *Structural Inorganic Chemistry*, Clarendon Press, Oxford (1975).

Up to now, we have described ideal crystals which are obtained by combining a basis with an infinite space lattice. We have also described at the other extreme noncrystalline structures which, except for local coordination, have no regularity or periodicity whatsoever in the atomic arrangement. In this chapter, we shall consider real crystals. First of all, we note that real crystals are finite in extent. Therefore, they have a surface as their boundary, where some of the atomic bonds are broken. The surface itself is then an imperfection. In addition, there are other occasional disruptions in periodicity within a crystal. Volumewise, the disrupted regions may be as small as 0.01% of the total volume. If this is so, the beginner starts to wonder why at all then we are bothered about such infrequent disruptions. This view would be fully justified if we were studying structure-insensitive properties such as the density of a crystal. If 0.01% of the atoms were missing from the atomic sites, the density of the crystal would be lower by the same amount (0.01%), which is small enough to be ignored for all practical purposes. However, many of the properties of crystalline materials that we deal with in engineering practice are structure-sensitive properties. They vitally depend on the presence or absence of imperfections, however infrequent they may be when present. For example, a few parts per million of aluminium impurity may radically change the character of a silicon semiconductor. If the impurity gets in inadvertently, it could destroy the rectifying or transistor action of junctions altogether. Likewise, line imperfections known as dislocations decrease the mechanical strength of crystals drastically.

Crystalline imperfections can be classified on the basis of their geometry as follows:

1. Point imperfections
2. Line imperfections
3. Surface imperfections
4. Volume imperfections.

We will discuss only the first three in this chapter. Volume imperfections can be foreign-particle inclusions, large voids or pores, or noncrystalline regions which have the dimensions of at least a few tens of Å.

Units

Quantity	SI units		Other units
	Unit	Symbol	
Concentration of point imperfections n/N	—		—
Enthalpy of formation of point imperfections ΔH_f	kilojoule per mole	kJ mol^{-1}	eV/entity
Dislocation density ρ	metre per cubic metre or per square metre	m m^{-3} or m^{-2}	in-in^{-3} or in^{-2}
Dislocation energy E	joule per metre	J m^{-1}	erg/cm
Shear modulus μ	giganewton per square metre	GN m^{-2}	dyne/cm ² , psi
Energy of surface imperfections γ	joule per square metre	J m^{-2}	erg/cm ²
Surface tension	newton per metre	N m^{-1}	dyne/cm

Constants

Gas constant $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$
--

6.1 Point Imperfections

Point imperfections are also referred to as zero-dimensional imperfections. As the name implies, they are imperfect point-like regions in the crystal. One or two atomic diameters is the typical size of a point imperfection. Different kinds of point imperfections are now described.

A *vacancy* refers to an atomic site from where the atom is missing, as shown in Fig. 6.1a. Figure 6.1d depicts a vacancy in platinum photographed in a field ion microscope.

We have already introduced the substitutional and the interstitial solid solutions in Chap. 5. A *substitutional impurity* (or solute) is a point imperfection. It refers to a foreign atom that substitutes for or replaces a parent atom in the crystal, see Fig. 6.1b. Aluminium and phosphorus doped in silicon are substitutional impurities in the crystal. An *interstitial impurity* is also a point imperfection. It is a small sized atom occupying the void space in the parent crystal, without dislodging any of the parent atoms from their sites, as shown in Fig. 6.1c. An atom can enter the interstitial void space only when it is substantially smaller than the parent atom. In close packed structures, the largest atom that can fit the octahedral and the tetrahedral voids have radii $0.414r$ and $0.225r$, respectively, where r is the radius of the parent atom. For example, carbon is an interstitial solute in iron. It occupies the octahedral voids in the high temperature FCC form of iron. The iron atom in the FCC crystal has a

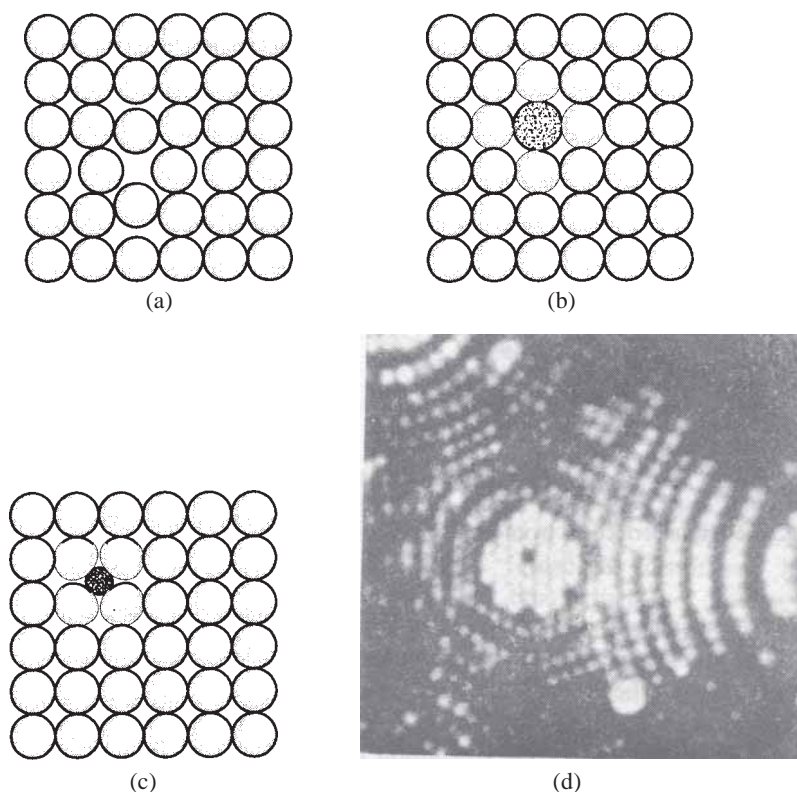


Fig. 6.1 Point imperfections in an elemental crystal: (a) vacancy; (b) substitutional impurity; (c) interstitial impurity; and (d) field ion micrograph of platinum showing a vacancy. [(d) *Courtesy: E.W. Mueller.*]

radius of 1.29 \AA , whereas the carbon atom has a radius of 0.71 \AA (covalent radius in graphite). The carbon radius is clearly larger than $0.414 \times 1.29 = 0.53 \text{ \AA}$, which is the size of the octahedral void. Therefore, there are strains around the carbon atom in the FCC iron, and the solubility is limited to 2 wt.%. In the room temperature BCC form of iron, the voids are still smaller and hence the solubility of carbon is very limited, that is, only 0.008 wt.%.

In ionic crystals, the formation of point imperfections is subject to the requirement that the *overall electrical neutrality is maintained*. An ion displaced from a regular site to an interstitial site is called a *Frenkel imperfection*, see Fig. 6.2a. As cations are generally the smaller ions, it is possible for them to get displaced into the void space. Anions do not get displaced like this, as the void space is just too small for their size. A Frenkel imperfection does not change the overall electrical neutrality of the crystal. The point imperfections in silver halides and CaF_2 are of the Frenkel type.

A pair of one cation and one anion can be missing from an ionic crystal as shown in Fig. 6.2b. The valency of the missing pair of ions should be equal to maintain electrical neutrality. Such a pair of vacant ion sites is called a *Schottky imperfection*. This type is dominant in alkali halides.

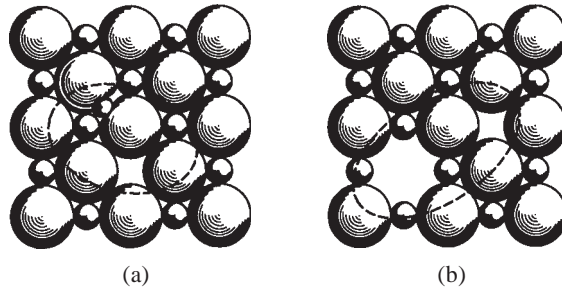


Fig. 6.2 Point imperfections in an ionic crystal: (a) Frenkel defect; and (b) Schottky defect.

We noted in Chap. 5 that trivalent cations such as Fe^{3+} and Cr^{3+} can substitute for trivalent parent Al^{3+} cations in the Al_2O_3 crystal. If, however, the valency of the substitutional impurity is not equal to the parent cation, additional point defects may be created due to such substitution. For example, a divalent cation such as Cd^{2+} substituting for a univalent parent ion such as Na^+ will, at the same time, create a vacant cation site in the crystal so that electrical neutrality is maintained.

Defect structures are produced when the composition of an ionic crystal does not correspond to the exact stoichiometric formula. Such defect structures have an appreciable concentration of point imperfections. Consider *deviations from the stoichiometric formula* in compounds of ZnO and FeO . In Zn_yO , where $y > 1$, the excess cations occupy interstitial voids. Such a compound can be produced by heating a stoichiometric compound in zinc vapour. The two electrons released from each zinc atom that enters the crystal stays around an interstitial cation, as shown in Fig. 6.3a. In Fe_xO , where $x < 1$, vacant cation sites are present. Such a compound can be produced by heating a stoichiometric FeO in an oxygen atmosphere. The two electrons required by each excess oxygen atom is donated by two Fe^{2+} ions which become Fe^{3+} (ferric) ions, see Fig. 6.3b.

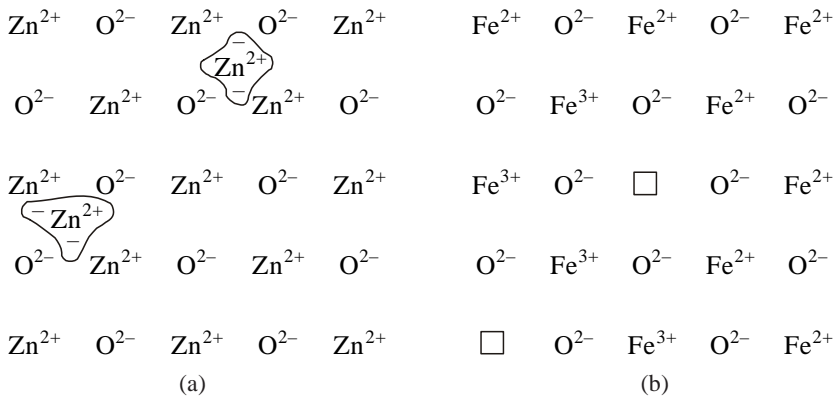


Fig. 6.3 Defects present in ZnO and FeO owing to deviations from stoichiometry. Vacant cation sites are indicated by \square .

The presence of a point imperfection introduces distortions in the crystal. If the imperfection is a vacancy, the bonds that the missing atom would have formed with its neighbours are not there. In the case of an impurity atom, as a result of the size difference, *elastic strains* are created in the region of the crystal immediately surrounding the impurity atom. The elastic strains are present irrespective of whether the impurity atom is larger or smaller than the parent atom. A larger atom introduces compressive stresses and corresponding strains around it, while a smaller atom creates a tensile stress-strain field. Similarly, an interstitial atom produces strains around the void it is occupying. All these factors tend to increase the enthalpy (or the potential energy) of the crystal. The work required to be done for creating a point imperfection is called the *enthalpy of formation* (ΔH_f) of the point imperfection. It is expressed in kJ mol^{-1} or $\text{eV/point imperfection}$. The enthalpy of formation of vacancies in a few crystals is shown in Table 6.1.

TABLE 6.1
Enthalpy of Formation of Vacancies in Some Crystals

Crystal	Kr	Cd	Pb	Zn	Mg	Al	Ag	Cu	Ni
ΔH_f kJ mol^{-1}	7.7	38	48	49	56	68	106	120	168
eV/vacancy	0.08	0.39	0.50	0.51	0.58	0.70	1.10	1.24	1.74

By virtue of the fact that a *point imperfection is distinguishable from the parent atom*, the configurational entropy of a crystal increases from zero for a perfect crystal to positive values with increasing concentration of the point imperfection, refer to Eqs. (2.5) and (2.12). When we introduce n point imperfections, in one mole of a crystal, the change in the free energy ΔG of the crystal can be written as

$$\Delta G = \Delta H - T\Delta S = n\Delta H_f - kT [N \ln N - (N - n) \ln (N - n) - n \ln n] \quad (6.1)$$

where N is Avogadro's number. The equilibrium state of the crystal will correspond to the minimum in its free energy, as shown in Fig. 6.4. The

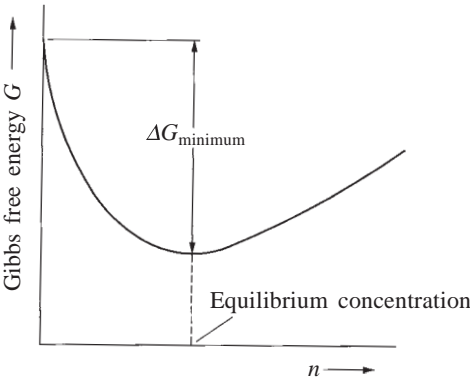


Fig. 6.4 The variation of Gibbs free energy G with the number of point imperfections n (schematic).

minimum in G also corresponds to the minimum in ΔG (note that ΔG is negative here). By setting $d\Delta G/dn = 0$, we obtain the equilibrium concentration:

$$\frac{n}{N} = \exp \left(-\frac{\Delta H_f}{RT} \right) \quad (6.2)$$

Example 6.1 Find the equilibrium concentration of vacancies in aluminium and nickel at 0 K, 300 K and 900 K.

Solution Using the data of Table 6.1 in Eq. (6.2), we obtain n/N . For aluminium, (i) at 0 K,

$$\frac{n}{N} = \exp \left(-\frac{68 \times 10^3}{8.314 \times 0} \right) = \exp (-\infty) = 0$$

(ii) at 300 K,

$$\begin{aligned} \frac{n}{N} &= \exp \left(-\frac{68 \times 10^3}{8.314 \times 300} \right) = \exp (-27.26) \\ &= 1.45 \times 10^{-12} \end{aligned}$$

Similarly, the other required values are obtained.

Temperature	0 K	300 K	900 K
Aluminium	0	1.45×10^{-12}	1.12×10^{-4}
Nickel	0	5.59×10^{-30}	1.78×10^{-10}

As ΔH_f is smaller for aluminium, the n/N values are larger for aluminium at all temperatures above 0 K.

In ionic crystals, point defects occur as pairs, the Frenkel defect being a cation vacancy and a cation interstitial, and the Schottky defect being a pair of cation and anion vacancies. In a stoichiometric compound, the concentration of Schottky imperfections in thermal equilibrium is given by

$$n^{\text{Sch}} = N \exp \left(-\frac{\Delta H_f^{\text{Sch}}}{2RT} \right) \quad (6.3)$$

where ΔH_f^{Sch} is the enthalpy of formation of one mole each of cation and anion vacancies. A similar equation describes the concentration of Frenkel defects.

Point imperfections of different types can interact with one another and lower the total energy. For example, a solute atom which is larger than the parent atom can have a smaller distortional energy if it stays close to a vacancy. This reduction in energy is called the *binding energy* between the two point imperfections. It is typically in the range 10–20% of the enthalpy of formation of point imperfections.

6.2 The Geometry of Dislocations

Line imperfections are called *dislocations*. Note that the word ‘dislocations’ is used by convention to denote only line imperfections, even if the word means any general discontinuity in the crystal. Line imperfections are one-dimensional imperfections in the geometrical sense. The concept of a dislocation was first put forward in 1930s. Abundant experimental evidence for the presence of dislocations in crystals has since accumulated. The static and dynamic properties of dislocations have been studied in great detail because of the vital role they play in determining the structure-sensitive properties of crystalline materials.

Dislocations are best understood by referring to two limiting straight-line types: (i) the edge dislocation, and (ii) the screw dislocation.

Discussing the *edge dislocation* first, Fig. 6.5a shows a perfect crystal, the top sketch depicting a three-dimensional view and the bottom one showing the

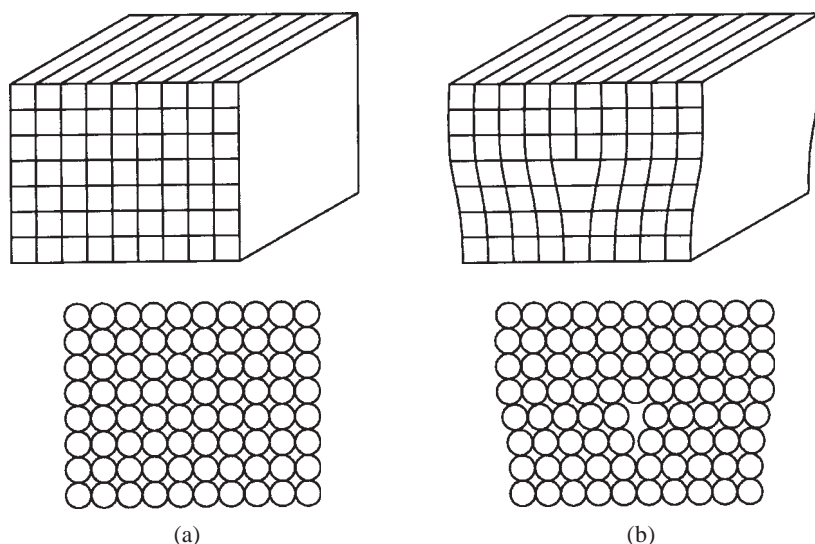


Fig. 6.5 An incomplete plane in a crystal results in an edge dislocation.

atoms on the front face. We can consider the perfect crystal to be made up of vertical planes parallel to one another and to the side faces. If one of these vertical planes does not extend from the top to the bottom of the crystal but ends part way within the crystal, as in Fig. 6.5b, a dislocation is present. In the lower sketch, notice the atomic arrangements on the front face. In the perfect crystal, the atoms are in equilibrium positions and all the bond lengths are of the equilibrium value. In the imperfect crystal on the right, just above the edge of the incomplete plane, the atoms are squeezed together and are in a *state of compression*. The bond lengths have been compressed to smaller than the equilibrium value. Just below the edge, the atoms are pulled apart and are in a

state of tension. The bond lengths have been stretched to above the normal value. This distorted configuration extends all along the edge into the crystal. Recall from Fig. 4.4 that the potential energy increases from the minimum value for the normal bond length (r_0) to higher values on both sides, that is, for an increase as well as a decrease in the bond length. Thus there is extra energy due to the distortion in the region immediately surrounding the edge of the incomplete plane. As the region of maximum distortion is centred around the edge of the incomplete plane, this distortion represents a line imperfection and is called an edge dislocation. Edge dislocations are symbolically represented by \perp or \top , depending on whether the incomplete plane starts from the top or from the bottom of the crystal. These two configurations are referred to as *positive* and *negative* edge dislocations.

An alternative way of looking at an edge dislocation is illustrated in Fig. 6.6. The hatched region $ABFE$ is part of the plane marked $ABCD$ in Fig. 6.6a. In Fig. 6.6b, the top part of the crystal above the hatched area $ABFE$

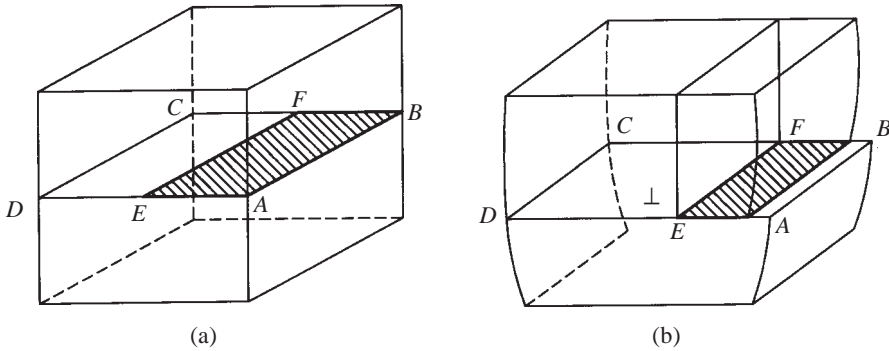


Fig. 6.6 Displacement of the top part with respect to the bottom by one step across the hatched area introduces an edge dislocation EF in the crystal.

is displaced to the left by one interatomic distance with respect to the bottom part of the crystal. No such displacement is done over the unhatched area $EFCD$. Starting from the right side face, the first vertical plane over the hatched area is now in alignment with the second vertical plane below the hatched area. The second plane above is in alignment with the third plane below and so on, till we reach the boundary EF of the hatched area. As the displacement operation ends here, there is no change in the alignment of the top and the bottom part of the vertical planes in the unhatched part. The displacement has thus resulted in an incomplete vertical plane over the edge EF , which thus becomes an edge dislocation.

The plane $ABCD$ is called a *slip plane* in dislocation terminology. The region $ABFE$ over which the displacement is done is called the slipped part of the slip plane. The region $EFCD$ over which the displacement is not done is called the unslipped part of the slip plane. The dislocation line EF is then

defined as *the boundary between the slipped and the unslipped parts* of the slip plane. This definition of a dislocation is, in fact, general and applies to any dislocation line.

In the above example, we considered a particular case in which the displacement was one interatomic distance. The magnitude and the direction of the displacement are defined by a vector called the *Burgers vector* (BV), which characterizes a dislocation line. The Burgers vector of a dislocation is determined as follows: Consider the perfect crystal shown in Fig. 6.7a. Starting from the point P , if we go up by x number of steps ($x = 4$ in Fig. 6.7a), then

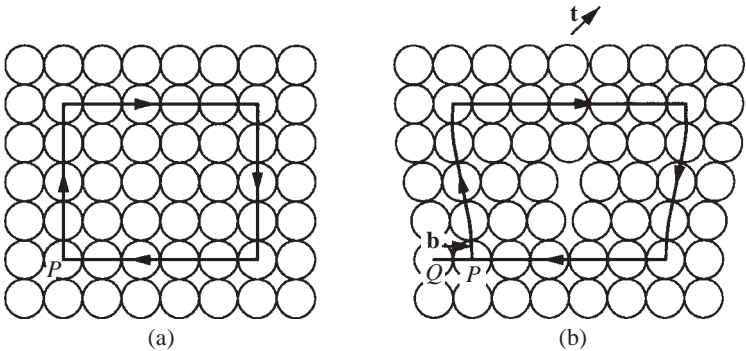


Fig. 6.7 Burgers circuits (a) in a perfect crystal; and (b) in an imperfect crystal with an edge dislocation.

take y steps to the right ($y = 5$ in the figure), then x steps down and finally y steps to the left, we end up at the starting point. We have now traced a *Burgers circuit* in taking these steps. We returned to the starting point because the region enclosed by the Burgers circuit is perfect, with no line imperfection cutting across it. If we now do the same operation on a crystal which has a dislocation in it as shown in Fig. 6.7b, starting from point P , we end up at Q . We need an extra step to return to point P (or to close the Burgers circuit). The magnitude and the direction of this step define the Burgers vector:

$$\text{BV} = \overrightarrow{QP} = \mathbf{b} \tag{6.4}$$

The Burgers vector is perpendicular to the edge dislocation line.

We have taken the Burgers circuit to be clockwise in the above example. The direction of the Burgers vector depends on the direction of the circuit, which can be clockwise or anticlockwise. To avoid this ambiguity, a unit vector \mathbf{t} is first assigned to denote the direction of the dislocation line. The direction vector is drawn tangential to the dislocation line at the point of interest. Then a right-hand screw (RHS) convention is followed in tracing the circuit. If we place the end of an ordinary (right handed) screw on the paper and turn the head clockwise, the screw tends to move into the plane of the paper. If the vector \mathbf{t} has a direction vertically into the plane of the paper as in Fig. 6.7b, using the RHS convention, the Burgers circuit should be drawn clockwise.

The other limiting type of dislocation is a *screw dislocation*. Consider the hatched area $AEFD$ on the plane $ABCD$ in Fig. 6.8a. As before, let the top part of the crystal over the hatched area be displaced by one interatomic distance to the left with respect to the bottom part, as shown in Fig. 6.8b. As a different area is hatched here as compared to the edge dislocation case in Fig. 6.6, for the

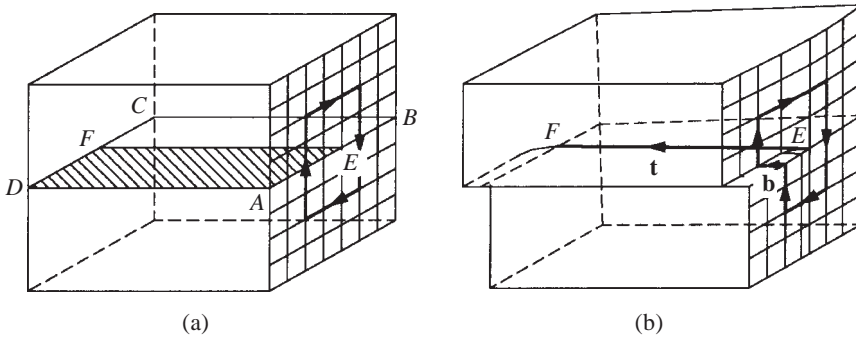


Fig. 6.8 A screw dislocation EF is created by displacement of the top part of the crystal with respect to the bottom across the hatched area.

same sense of displacement, we now produce a screw dislocation at the boundary EF between the displaced and the undisplaced parts of the slip plane. Let the \mathbf{t} vector be defined such that it points from right to left, as shown in Fig. 6.8b. The Burgers circuit is then drawn using the RHS convention on the right side face in a clockwise sense. The Burgers vector is determined by the step needed to close the circuit. In this case of a screw dislocation, the Burgers vector \mathbf{b} has the same direction as the \mathbf{t} vector. *The Burgers vector is parallel to the screw dislocation line.*

We note that there is no extra plane at the screw dislocation in contrast to what we saw in the edge dislocation. Examine the atomic arrangement around the dislocation line. If we go round the dislocation line once as illustrated in Fig. 6.9, we move along the line by a step equal in magnitude to the Burgers vector. If we turn a screw through one full rotation, we move the screw backward or forward by a distance equal to its pitch. This feature has led to the name *screw dislocation*. The atomic bonds in the region immediately surrounding the dislocation line have undergone a shear distortion, as seen clearly in Fig. 6.9. The vertical atomic planes parallel to the side faces in the perfect crystal of Fig. 6.8a have become continuous due to the displacement in Fig. 6.8b. They can be compared to a spiral staircase, whose central pillar coincides with the screw dislocation line. Screw dislocations are symbolically represented by \odot or \otimes , depending on whether the Burgers vector and the \mathbf{t} vector are parallel or antiparallel. These two cases are referred to as *positive* and *negative* screw dislocations.

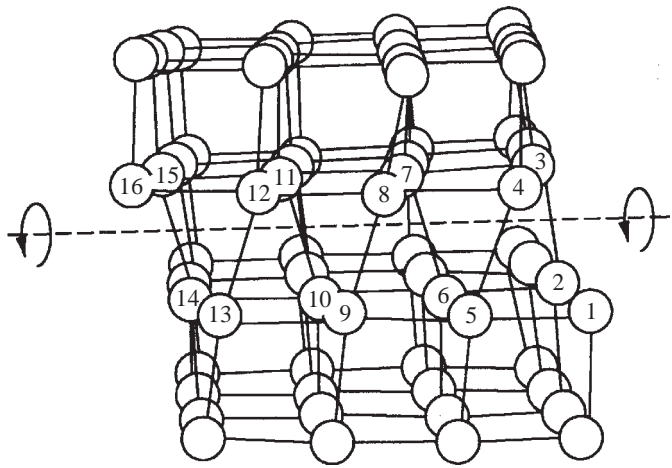


Fig. 6.9 The atoms around a screw dislocation line are arranged in a spiral screw-like fashion.

Any general dislocation line is a combination of the edge and the screw types. In Fig. 6.10, the hatched area on the slip plane is bounded by a curved line. If we do the same displacement over the hatched area as before, the

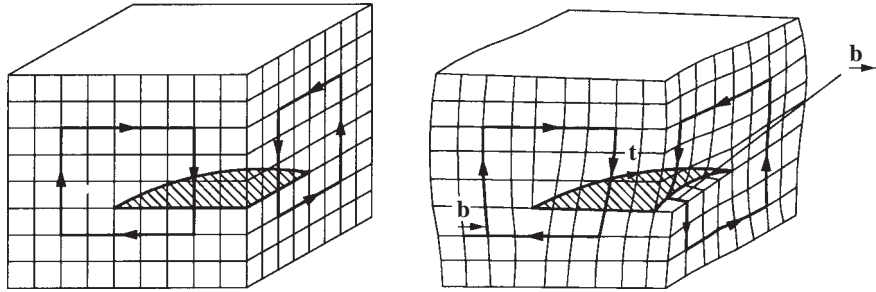


Fig. 6.10 Displacement across the hatched area with a curved boundary produces a mixed dislocation line at the boundary.

boundary of the hatched area that lies within the crystal becomes a curved dislocation line. The dislocation has pure screw character at the right face and pure edge character on the front face. It has mixed character in between. Note that the \mathbf{t} vector goes into the plane of the paper on the front face. It changes continuously its direction as the dislocation is curved and finally it emerges out on the right side face. Burgers circuits are drawn on both the front face and the right side face, following the RHS convention. The circuit is clockwise on the front face as the \mathbf{t} vector is facing inwards. It is anticlockwise on the right side face as the \mathbf{t} vector is emerging out of that face. The Burgers vector \mathbf{b} has the same magnitude and direction in both cases. Thus we obtain the important result

that the *Burgers vector is invariant*, that is, it has the same direction and magnitude all along a dislocation line, irrespective of the character of the dislocation (edge, screw or mixed).

In Fig. 6.11a, the hatched area on the slip plane is enclosed by an elliptical loop. When the displacement is carried out as before over the hatched area, an

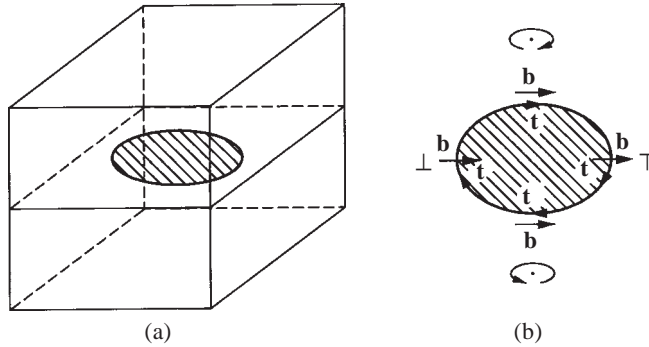


Fig. 6.11 Displacement over the hatched area enclosed by an elliptical loop produces a dislocation loop with varying character.

elliptical dislocation loop is produced at the boundary. In Fig. 6.11b, a plan view of the loop is shown. It has four points on it, corresponding to pure positive edge \perp , pure positive screw \curvearrowright , pure negative edge T and pure negative screw \curvearrowleft . In between these points, the dislocation has a mixed character. Here, the Burgers vector can be resolved in two directions, parallel and perpendicular to the line, thereby defining the screw and the edge components of the combination.

Dislocations have certain other geometrical characteristics. For instance, the vectorial sum of the Burgers vectors of the dislocations meeting at a point called the *node* must be zero, analogous to Kirchhoff's law for electrical currents, meeting at a junction. For consistency, the \mathbf{t} vectors of all the dislocations meeting at a node must either point towards it or away from it. A dislocation line cannot end abruptly within the crystal. It either ends at a node or at the surface. Alternatively, it can close on itself as a loop.

6.3 Other Properties of Dislocations

Dislocations have distortional energy associated with them as is evident from the compressive and the tensile strains around an edge dislocation or the shear strains around a screw dislocation. As a first approximation, we take these strains to be elastic strains. The elastic strain energy E per unit length of a dislocation of Burgers vector \mathbf{b} is approximately given by

$$E \cong \mu b^2/2 \quad (6.5)$$

where μ is the shear modulus of the crystal.

Example 6.2 Compute the line energy of dislocations in BCC iron. The Burgers vector in iron is of the $1/2 \langle 111 \rangle$ type. The shear modulus of iron is 80.2 GN m^{-2} .

Solution From the lattice parameter of BCC iron ($a = 2.87 \text{ \AA}$), we obtain the magnitude of the Burgers vector $\mathbf{b} = 2.87\sqrt{3}/2 = 2.49 \text{ \AA}$. The line energy of the dislocation from Eq. (6.5) is

$$\begin{aligned} E &= 80.2 \times 10^9 \times 2.49^2 \times 10^{-20}/2 \\ &= 2.49 \times 10^{-9} \text{ J m}^{-1} \end{aligned}$$

Dislocations in real crystals can be classified as full and partial dislocations. For a *partial dislocation*, the Burgers vector is a fraction of a lattice translation. For a full dislocation, the Burgers vector is an integral multiple of a lattice translation. We will deal with full dislocations only. As *the elastic strain energy of a dislocation is proportional to the square of the Burgers vector* equation (6.5), dislocations tend to have as small a Burgers vector as possible. Consider, for example, an edge dislocation in a simple cubic crystal, where the unit translation vector is \mathbf{b} . Let two incomplete planes be together on the slip plane so that the Burgers vector of this dislocation is $2\mathbf{b}$. The two planes will have a tendency to break up as two separate planes, that is, the giant dislocation will tend to break up into two dislocations of Burgers vector \mathbf{b} each.

Dislocation reaction: $2\mathbf{b} \rightarrow \mathbf{b} + \mathbf{b}$
 Energy change: $\mu(2b)^2/2 \rightarrow \mu b^2/2 + \mu b^2/2$

The elastic energy decreases by 50% during this breakup. In general, a dislocation with Burgers vector \mathbf{b}_1 will tend to dissociate into two (or more) dislocations of Burgers vectors \mathbf{b}_2 and \mathbf{b}_3 if

$$b_1^2 > b_2^2 + b_3^2 \tag{6.6}$$

For the same reason, full dislocations tend to have the smallest lattice translation as the Burgers vector. The most probable Burgers vectors of full dislocations in cubic crystals are given in Table 6.2.

TABLE 6.2
Burgers Vectors of Dislocations in Cubic Crystals

Monoatomic FCC	$1/2 \langle 110 \rangle$
Monoatomic BCC	$1/2 \langle 111 \rangle$
Monoatomic SC	$\langle 100 \rangle$
NaCl structure	$1/2 \langle 110 \rangle$
CsCl structure	$\langle 100 \rangle$
DC structure	$1/2 \langle 110 \rangle$

Note that the Burgers vector in a CsCl crystal cannot be from a body corner to the body centre, as this is not a full lattice translation. If we have an edge dislocation with Burgers vector of $\frac{1}{2} \langle 111 \rangle$ in a CsCl crystal, the extra plane will have excess electrical charges. The crystal as a whole will not be electrically neutral and, therefore, such a configuration is not stable. Similarly, the Burgers vector in a NaCl crystal cannot be from the centre of a chlorine ion at the cube corner to the centre of a sodium ion halfway along the cube edge. This restriction in ionic crystals tends to make the Burgers vectors large. For example, in the FCC crystal of copper, the Burgers vector of a full dislocation is only 2.55 Å in magnitude, whereas in the NaCl crystal, the Burgers vector is 3.95 Å.

The distortional geometry above the slip plane for a positive edge dislocation is opposite in sense to the distortion above the slip plane for a negative edge dislocation. The former is a region of compressive strains, while the latter is a region of tensile strains. Therefore, when two edge dislocations of opposite sign are in the same slip plane, they tend to attract each other. They may totally annihilate each other, that is, the distortional strain fields may superpose and exactly cancel out both above and below the slip plane if the Burgers vectors are also of the same magnitude. Dislocations of the same sign, on the other hand, repel each other. These characteristics are also true of screw dislocations.

Dislocations can interact with point imperfections if there is a decrease in the total energy (the energy of the dislocation plus the energy of the point imperfection). A substitutional solute atom is different in size from the parent atom and hence causes a local tensile or compressive strain field around it in the crystal. Such a strain centre can elastically interact with an edge dislocation which has compressive and tensile strains on either side of the slip plane. A smaller substitutional atom can reduce the total energy by displacing a parent atom in the compressive region. A larger atom will displace a parent atom in the tensile region. Similarly, interstitial solutes which are usually too big to fit into the interstitial voids can enter the more openly packed central core region of an edge dislocation. Such an interaction is particularly strong in mild steel, where interstitial carbon atoms form atmospheres around dislocations in the BCC crystal of iron. This interaction has an important effect on the mechanical behaviour of mild steel.

Dislocations are usually present in a crystal, as a result of *accidents during growth* of the crystal from the melt or as a result of prior mechanical deformation of the crystal. Unlike point imperfections, they are not thermodynamically stable, as the enthalpy of the crystal increases much more rapidly with their presence than the entropy. Many of them can be removed by heating the crystal to high temperatures, where the thermal energy will allow them to mutually cancel each other or to move out of the crystal through the surface. The density of dislocations in a crystal is measured by counting the number of points at which they intersect a random cross-section of the crystal. These points, called *etch pits*, can be seen under the microscope, after chemical etching of a prepolished surface. In an annealed crystal, the dislocation density ρ is in the range of 10^8 – 10^{10} m^{-2} . The dislocation density can also be expressed as the total length of dislocation lines in unit volume, m m^{-3} . In order of magnitude, these two densities are the same.

In Fig. 6.6, if the area of the hatched region $ABFE$ had been gradually increased, keeping EF straight and parallel to the side faces, the edge dislocation defined by the boundary EF would have successively occupied positions to the left starting from the right end. In other words, the dislocation would have *moved* from right to left. A dislocation already present in a crystal can be made to move in this manner by an externally applied stress. Figure 6.12 shows a

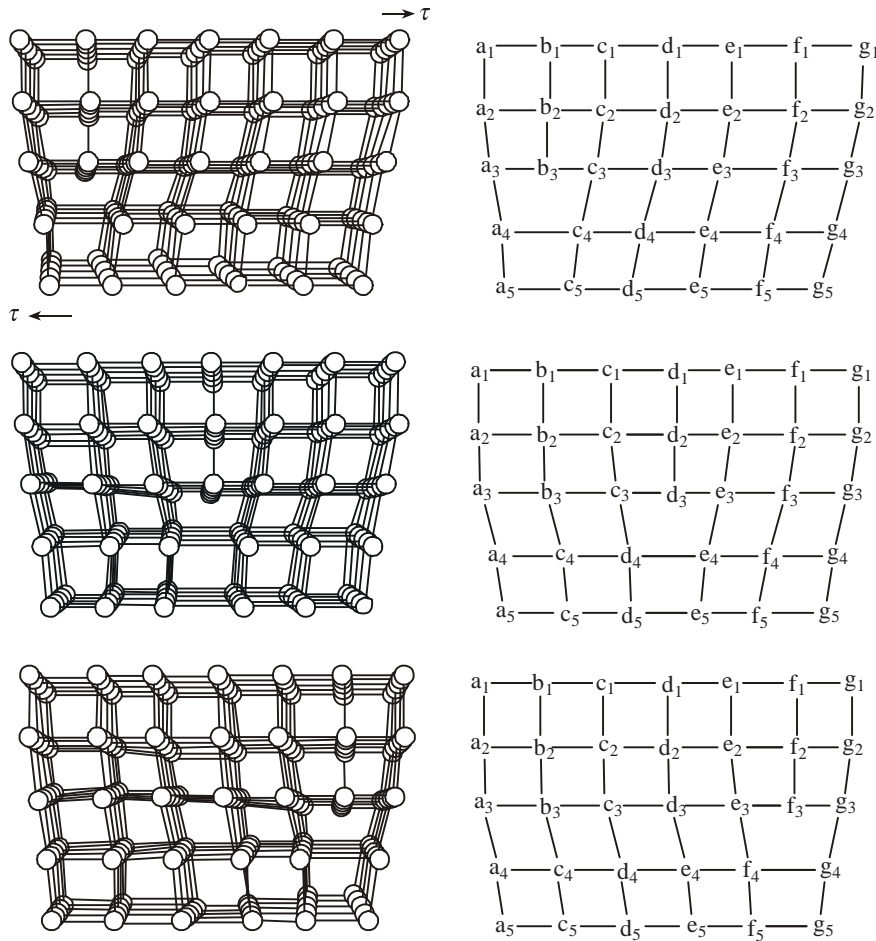


Fig. 6.12 Successive positions of an edge dislocation as it moves in response to an externally applied shear stress τ .

shear stress τ being applied parallel to the slip plane which has an edge dislocation lying on it. If the applied stress is sufficiently large, *the dislocation will move in such a direction as to bring about the displacement corresponding to the sense of the shear stress*. The sense of the stress in Fig. 6.12 is such that the top of the crystal will be displaced to the right in relation to the bottom part. This displacement can occur with reference to the slip plane by the movement

towards the right of the (positive) edge dislocation lying on it. The successive positions of the dislocation during this motion are illustrated in Fig. 6.12. The atoms have been labelled to show their positions before and after the dislocation motion. It is important to note that there is *no bodily shift* of the incomplete plane during the motion and that the configuration of the dislocation is what is moving. The configuration is shifted from one position to the next by *small variations of bond lengths* in the dislocation region.

The dislocation can eventually reach the surface in Fig. 6.12. As soon as this happens, the compressive and the tensile strains will be relieved at the free surface, and we would have a step on the right side face. This step at the surface does not have the compressive and tensile strains characteristic of the edge dislocation. It should not be called a dislocation any more. We say that the dislocation disappears on reaching the surface, leaving a step behind. The magnitude of the step is equal to the Burgers vector of the dislocation.

The motion of a dislocation on a plane that contains the direction vector \mathbf{t} and the Burgers vector \mathbf{b} is called the *glide motion*. The motion of the edge dislocation that we considered with reference to Fig. 6.12 is glide motion. The dislocation loop in Fig. 6.11 can glide on the slip plane as \mathbf{b} and \mathbf{t} lie on this plane at all points around the loop. In fact, it cannot glide on any other plane. As there is only one plane which can contain both \mathbf{b} and \mathbf{t} when they are not parallel, the glide plane is uniquely defined for a pure edge dislocation as well as for a mixed dislocation. The unit normal defining the glide plane is given by the cross product of \mathbf{b} and \mathbf{t} . For a screw dislocation, \mathbf{b} and \mathbf{t} are parallel or antiparallel so that a glide plane is not uniquely defined. The cross product of \mathbf{b} and \mathbf{t} is zero for a screw dislocation.

Example 6.3 A circular dislocation loop has edge character all round the loop. What is the surface on which this dislocation can glide?

Solution By definition, the Burgers vector is perpendicular to an edge dislocation line. Also, the Burgers vector is invariant. The edge dislocation can glide only on a surface that contains both the Burgers vector and the \mathbf{t} vector. These considerations are satisfied only when the given dislocation moves on a cylindrical surface containing the loop.

The motion of an edge dislocation on a plane perpendicular to the glide plane is called *climb motion*. As the edge dislocation moves above or below the slip plane in a perpendicular direction, the incomplete plane either shrinks or increases in extent. This kind of shifting of the edge of the incomplete plane is possible only by *subtracting or adding rows of atoms* to the extra plane. Climb motion is said to be *nonconservative*. This is in contrast to the glide motion which is conservative and does not require either addition or subtraction of atoms from the incomplete plane. During climbing up of an edge dislocation, the incomplete plane shrinks. Atoms move away from the incomplete plane to other parts of the crystal. During

climbing down, the incomplete plane increases in extent. Atoms move into the plane from other parts of the crystal. This results in an interesting interaction between the climb process and the vacancies in the crystal.

Example 6.4 If there are 10^{10} m^{-2} of edge dislocations in a simple cubic crystal, how much would each of these climb down on an average when the crystal is heated from 0 to 1000 K? The enthalpy of formation of vacancies is 100 kJ mol^{-1} . The lattice parameter is 2 \AA . The volume of one mole of the crystal is $5.5 \times 10^{-6} \text{ m}^3$ (5.5 cm^3).

Solution At equilibrium, there are no vacancies in the crystal at 0 K. To maintain equilibrium concentration, the number of vacancies that must be created on heating from 0 to 1000 K is given by

$$\begin{aligned} n &= N \exp(-\Delta H_f/RT) \\ &= 6.023 \times 10^{23} \exp[(-100 \times 1000)/(8.314 \times 1000)] \\ &= 3.60 \times 10^{18} \text{ mol}^{-1} \\ &= 6.54 \times 10^{23} \text{ m}^{-3} \end{aligned}$$

If the edge dislocations in the crystal climb down, atoms will be added to the extra plane and these atoms coming from the other parts of the crystal would create vacancies. As one step is 2 \AA , 5×10^9 atoms are required for 1 m of the dislocation line to climb down by one step. The average amount of climb down is then $6.54 \times 10^{23}/(5 \times 10^9 \times 10^{10}) = 1.31 \times 10^4$ steps $= 2.62 \times 10^{-6} \text{ m}$.

As there is no unique glide plane defined for a screw dislocation, it can change its slip plane during its motion. For example, the common slip planes in an FCC crystal are of $\{111\}$ type. Let a screw dislocation moving on a $(1\bar{1}1)$ plane come across some kind of an obstacle on this plane. If the shear stress is sufficient on a $(11\bar{1})$ plane, which is nonparallel to the $(1\bar{1}1)$ plane, then the screw dislocation can *cross-slip* into that plane and continue to glide. Figure 6.13 illustrates the cross-slip of a screw dislocation. As the glide plane

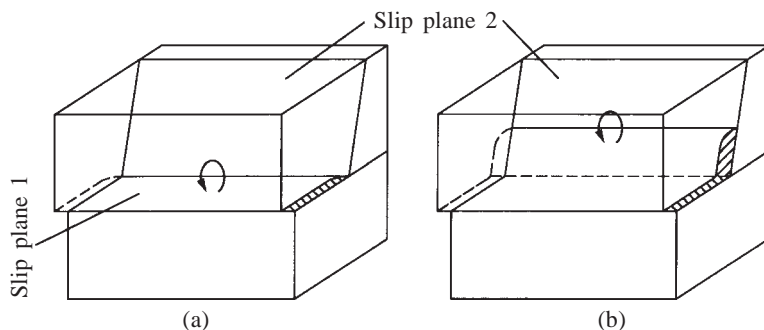


Fig. 6.13 A screw dislocation cross-slips from one slip plane onto another nonparallel slip plane (cont.).

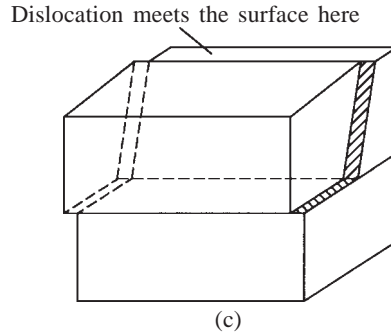


Fig. 6.13 A screw dislocation cross-slips from one slip plane onto another nonparallel slip plane.

(or the slip plane) is uniquely defined for an edge or a mixed dislocation, they cannot cross-slip in this manner.

6.4 Surface Imperfections

Surface imperfections are two dimensional in the mathematical sense. They refer to regions of distortions that lie about a surface having a thickness of a few atomic diameters. The external surface of a crystal is an imperfection in itself as the atomic bonds do not extend beyond the surface. External surfaces have surface energies that are related to the number of bonds broken at the surface. For example, consider a close packed plane as the surface of a close packed crystal. An atom on the surface of this crystal has six nearest bonding neighbours on the surface plane, three below it, and none above. Therefore, three out of twelve neighbours of an atom are missing at the surface. The surface energy of a crystal bears a relationship to this number.

Example 6.5 The surface of a copper crystal is of the $\{111\}$ type. Calculate the surface energy (enthalpy) of copper.

Solution

The bond energy per atom of copper

$$\begin{aligned}
 &= \text{bond energy per bond} \times \text{no. of bonds per atom} \times 1/2 \\
 &= (56.4 \times 1000 \times 12) / (6.023 \times 10^{23} \times 2) \\
 &= 5.62 \times 10^{-19} \text{ J}
 \end{aligned}$$

Three out of twelve bonds are broken at the surface.

$$\text{Energy of broken bonds per atom} = 5.62 \times 10^{-19} \times 1/4 \text{ J}$$

No. of atoms on $\{111\}$ planes in copper (see Problem 3.11)

$$= 1.77 \times 10^{19} \text{ m}^{-2}$$

$$\begin{aligned}
 \text{Surface enthalpy of copper} &= 5.62 \times 10^{-19} \times 1/4 \times 1.77 \times 10^{19} \\
 &= 2.49 \text{ J m}^{-2} \\
 &= 2490 \text{ erg/cm}^2.
 \end{aligned}$$

Experimentally measured surface free energies for some crystals are listed in Table 6.3.

TABLE 6.3
Surface Free Energies of Some Crystals (J m⁻²)

NaCl	LiF	CaF ₂	Ag	MgO	Si	Fe	Au	Cu
0.30	0.34	0.45	1.14	1.20	1.24	1.40	1.40	1.65

In addition to the external surface, crystals may have surface imperfections inside. A piece of iron or copper is usually not a single crystal. It consists of a number of crystals and is said to be *polycrystalline*. During solidification or during a process in the solid state called *recrystallization*, new crystals form in different parts of the material. They are randomly oriented with respect to one another. They grow by the addition of atoms from the adjacent regions and eventually impinge on each other. When two crystals impinge in this manner, the atoms caught in between the two are being pulled by each of the two crystals to join its own configuration. They can join neither crystal due to the opposing forces and, therefore, take up a compromise position. These positions at the boundary region between two crystals are so distorted and unrelated to one another that we can compare the boundary region to a noncrystalline material. The thickness of this region is only a few atomic diameters, because the opposing forces from neighbouring crystals are felt by the intervening atoms only at such short distances. The boundary region is called a *crystal boundary* or a *grain boundary* and is depicted in Fig. 6.14.

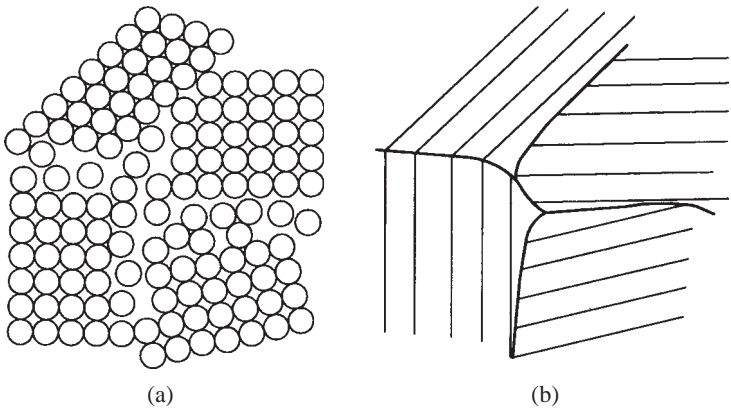


Fig. 6.14 The atomic arrangements at grain boundaries are distorted and unrelated.

The crystal orientation changes sharply at the grain boundary. The orientation difference is usually greater than $10\text{--}15^\circ$. For this reason, the grain boundaries are also known as *high angle boundaries*. The average number of nearest neighbours for an atom in the boundary of a close packed crystal is 11, as compared to 12 in the interior of the crystal. On an average, one bond out of the 12 bonds of an atom is broken at the boundary. The grain boundary between two crystals, which have different crystalline arrangements or differ in composition, is given a special name, viz. an *interphase boundary* or an *interface*. In an Fe-C alloy, the energies of grain boundaries and interfaces are compared in the following manner:

Grain boundary between BCC crystals	0.89 J m^{-2}
Grain boundary between FCC crystals	0.85 J m^{-2}
Interface between BCC and FCC crystals	0.63 J m^{-2}

When the orientation difference between two crystals is less than 10° , the distortion in the boundary is not so drastic as compared with a noncrystalline material. Such boundaries have a structure that can be described by means of arrays of dislocations. They are called low angle boundaries. Figure 6.15 shows a *low angle tilt boundary*, where neighbouring crystalline regions are tilted with respect to each other by only a small angle. The tilt boundary can be described

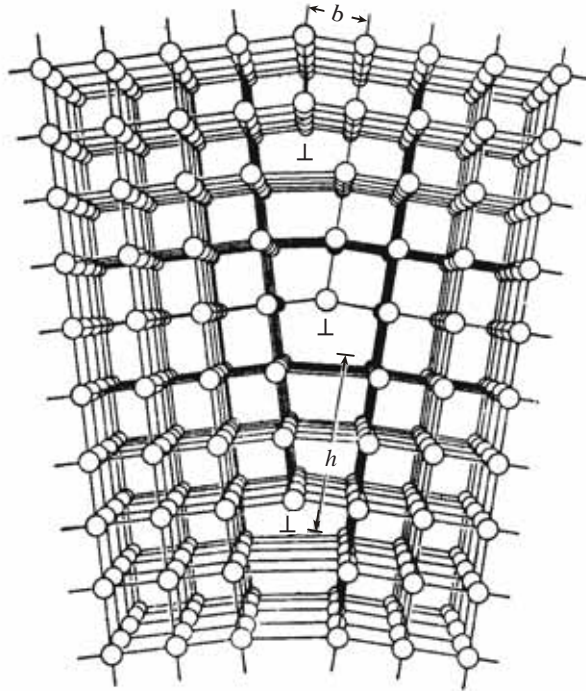


Fig. 6.15 A tilt boundary consists of equally spaced edge dislocations of the same sign one above the other.

as a set of parallel, equally-spaced edge dislocations of the same sign located one above the other. The angle of tilt θ is related to the Burgers vector b of the edge dislocations:

$$\frac{b}{h} = \tan \theta \quad (6.7)$$

where h is the vertical spacing between two neighbouring edge dislocations. For small angles of tilt, $b/h \sim \theta$. Similar low angle boundaries formed by screw dislocations are called *twist boundaries*. Note that both tilt and twist boundaries are planar surface imperfections in contrast to grain boundaries which need not be planar.

Stacking faults are also planar surface imperfections created by a fault (or error) in the stacking sequence of atomic planes in crystals. Consider the stacking arrangement in an FCC crystal

$$\downarrow$$

$$\dots ABCABCABCABC \dots$$

If an A plane indicated by an arrow above is missing, the stacking sequence becomes

$$\dots ABCBCBCABC \dots$$

The stacking in the missing region is $\dots BCBC \dots$ which is HCP stacking. This thin region is a surface imperfection and is called a stacking fault. The number of nearest neighbours in the faulted region remains 12 as in the perfect regions of the crystal, but the second nearest neighbour bonds in the faulted region are not of the correct type for the FCC crystal. Hence, a small surface energy is associated with the stacking fault, in the range $0.01\text{--}0.05 \text{ J m}^{-2}$ ($10\text{--}50 \text{ erg/cm}^2$). Similarly, we can define a stacking fault in an HCP crystal as a thin region of FCC stacking.

Another planar surface imperfection is a *twin boundary*. The atomic arrangement on one side of a twin boundary is a mirror reflection of the arrangement on the other side, as illustrated in Fig. 6.16. Twin boundaries occur

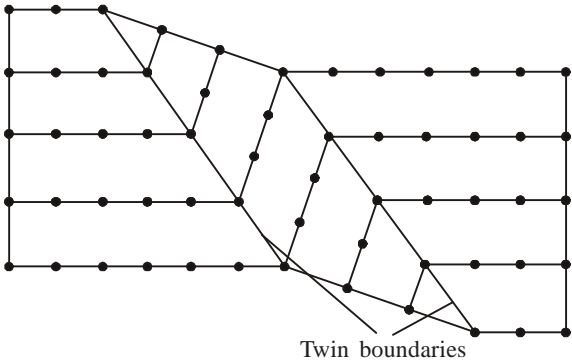


Fig. 6.16 The atomic arrangement on one side of a twin boundary is a mirror reflection of that on the other side.

in pairs such that the orientation change introduced by one boundary is restored by the other, as shown in Fig. 6.16. The region between the pair of boundaries is called the twinned region. Twin boundaries are easily identified under an optical microscope. Figure 6.17 shows the microstructure of brass with a number of twins in the grains. Twins which form during the process of recrystallization are called *annealing twins*, and those which form during plastic deformation of the material are called *deformation twins*.

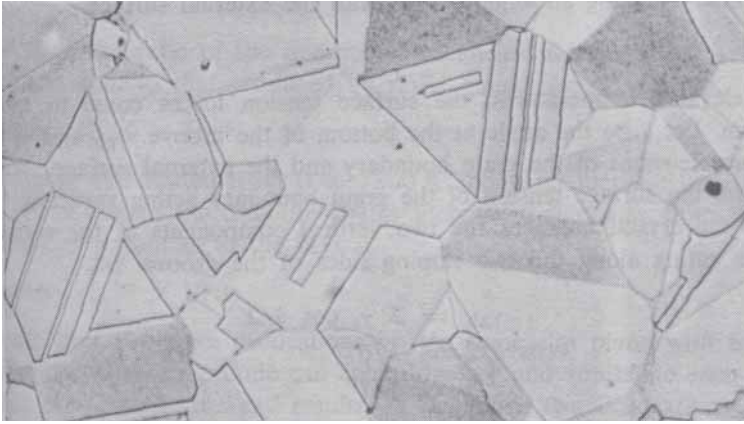


Fig. 6.17 Annealing twins and grain boundaries in brass (J. Nutting and R.G. Baker, *The Microstructure of Metals*, with permission from The Institute of Metals, London.)

Surface imperfections are not stable in a thermodynamic sense. They are present as metastable imperfections. If the thermal energy is increased by heating a crystal close to its melting point, many of the surface imperfections can be removed. The grain boundary area decreases as a polycrystalline material is heated above $0.5T_m$, where T_m is the melting point in K. Larger crystals grow at the expense of smaller crystals. Even though the average size of a crystal increases during this grain growth, the number of crystals decreases, resulting in a net decrease in the grain boundary area per unit volume of the material.

Example 6.6 A metal heated to elevated temperatures exhibits grooves on the surface at positions where the grain boundaries meet the surface, see Fig. 6.18. Assume that the ratio of the surface energy of the free surface to that of the grain boundary is 3 to 1. Compute the angle at the bottom of the groove of a boundary, which makes an angle of 90° with the external surface.

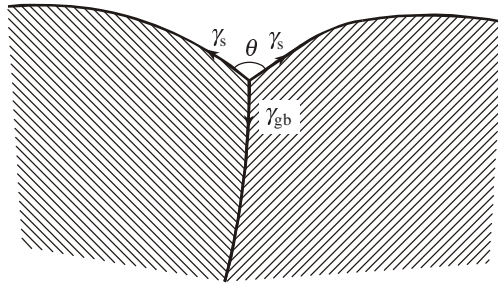


Fig. 6.18 Grooves are formed at high temperature on the surface along the lines where the grain boundaries meet the external surface.

Solution At elevated temperatures, the surface tension forces come to an equilibrium. Let θ be the angle at the bottom of the groove. γ_{gb} and γ_s are the surface tensions of the grain boundary and the external surface. At equilibrium, the surface tension of the grain boundary acting vertically inwards into the crystal balances the two vertical components of the surface tensions acting along the two sloping sides of the groove. So,

$$\gamma_{gb} = 2\gamma_s \cos\left(\frac{\theta}{2}\right)$$

$$\cos \frac{\theta}{2} = \frac{1}{6}$$

$$\theta = 161^\circ.$$

SUMMARY

1. Crystal imperfections can be classified according to their geometry.
2. Point imperfections are vacancies and impurity atoms in elemental crystals. They are Frenkel and Schottky imperfections in ionic crystals.
3. A certain concentration of point imperfections is present in equilibrium in a crystal as the configurational entropy due to their presence lowers the free energy of the crystal.
4. Edge and screw dislocations are the limiting types of line imperfections. Any general dislocation can be resolved into edge and screw components.
5. The elastic strain energy of dislocations varies as the square of their Burgers vector. This favours small Burgers vectors.

6. The glide motion of dislocations under an applied shear stress can deform a crystal. The climb of an edge dislocation occurs with the aid of point imperfections. Only screw dislocations can cross-slip from one slip plane to another nonparallel slip plane.
7. The external surface and the internal grain boundaries are typical nonplanar surface imperfections in a crystal.
8. The atomic arrangement can be described in detail at planar boundaries such as tilt boundaries, twin boundaries and stacking faults.

PROBLEMS

- 6.1** Calculate the ratio of the number of vacancies in equilibrium at 300 K in aluminium to that produced by rapid quenching from 800 K.
Answer: 3.99×10^{-8} .
- 6.2** A copper rod heated from 0 to 1300 K shows an increase in length of 2.1%. Estimate the fraction of this increase that comes from the formation of vacancies in the crystal.
Answer: 2.38×10^{-4} .
- 6.3** Using sketches, show how two edge dislocations on the same slip plane with Burgers vectors of equal magnitude can come together and annihilate each other.
- 6.4** What are the possible end results if the above two dislocations are in adjacent slip planes?
- 6.5** Make two neat sketches to show the climbing up and the climbing down of an edge dislocation. What happens to the vacancy concentration in the crystal during each process?
- 6.6** In a simple cubic crystal ($a = 3 \text{ \AA}$), a positive edge dislocation 1 mm long climbs down by $1 \text{ }\mu\text{m}$. How many vacancies are lost or created?
Answer: 1.1×10^{10} created.
- 6.7** Estimate approximately the atomic per cent of interstitial carbon required to fill all the core sites of edge dislocations in iron. Assume the edge dislocation density to be 10^{11} m^{-2} . Comment on the result that you obtain.
Answer: 5×10^{-9} .
- 6.8** Does the Burgers vector change with the size of the Burgers circuit? Explain.
- 6.9** Draw a Burgers circuit that encloses a positive and a negative edge dislocation each with one incomplete plane in a simple cubic crystal. What is the Burgers vector obtained? Comment on your result.

- 6.10** The Burgers vector of a mixed dislocation line is $1/2[110]$. The dislocation line lies along the $[112]$ direction. Find the slip plane on which this dislocation lies. Find also the screw and the edge components of the Burgers vector.

Answer: $(1\bar{1}0)$; $1/6 [112]$ and $1/3 [11\bar{1}]$.

- 6.11** Sketch the arrangement of atoms in Fig. 6.12, after the edge dislocation has reached the right side face.

- 6.12** Distinguish between the direction of the dislocation line, the Burgers vector, and the direction of motion for both edge and screw dislocations, differentiating between positive and negative types.

- 6.13** An aluminium crystal has a dislocation density of 10^{10} m^{-2} . The shear modulus of aluminium is 25.94 GN m^{-2} . Calculate the elastic energy of line imperfections stored in the crystal.

Answer: 10.61 J m^{-3} .

- 6.14** Show both graphically and analytically that the first two dislocations add to give the third dislocation in the following reaction:

$$1/6 [21\bar{1}] + 1/6 [121] \rightarrow 1/2 [110].$$

- 6.15** Calculate the square of the Burgers vectors of the above dislocations and determine whether the reaction from left to right is energetically favoured.

- 6.16** Calculate the surface enthalpy of copper when the external surface is of (i) $\{100\}$, and (ii) $\{110\}$ type.

Answer: 2.87 and 2.54 J m^{-2} .

- 6.17** A gold wire of radius of 0.08 mm is held at a temperature just below its melting point and is strained in tension by a small suspended weight. Under a load of 0.736 mN (including the weight of the wire itself), the wire maintains a constant length. Under smaller loads, it shortens; under larger loads, it lengthens. What is the surface tension of gold at this temperature?

Answer: 1.46 N m^{-1} .

- 6.18** The interfaces between three phases α , β , and γ meet along an edge. The angles subtended at the edge by the three phases are respectively 120° , 105° and 135° . If the surface energy of the α - β boundary is 1.00 J m^{-2} , find the surface energy of β - γ and γ - α interfaces.

Answer: 1.225 and 1.366 J m^{-2} .

- 6.19** The grains in a polycrystalline solid are space-filling polyhedra which can be approximated to spheres of equivalent volume. Calculate the change in the grain boundary energy per m^3 of the solid, when the average grain diameter increases from 0.01 mm to 0.1 mm . From the result obtained, deduce whether grain growth is an energetically favoured process.

Answer: A decrease of 90% .

- 6.20** Calculate the spacing between dislocations in a tilt boundary in FCC nickel, when the angle of tilt is 2° .

Answer: 71.3 Å.

- 6.21** At a twin boundary, the orientation difference between adjacent parts of the crystal is as large as that at a grain boundary. Yet, the twin boundary energies are in the range $0.01\text{--}0.05\text{ J m}^{-2}$, as compared to grain boundary energies in the range $0.2\text{--}0.6\text{ J m}^{-2}$. What is the reason for this difference?

- 6.22** In an aluminium crystal, the number of vacancies triples itself, on increasing the temperature from 300 K to 312.5 K. Calculate the enthalpy of formation of vacancies.

Answer: 68.5 kJ mol^{-1} .

- 6.23** Arrive at an order of magnitude value for the concentration of vacancies in any elemental crystal just below its melting point from the following data:

Element	ΔH_f , kJ mol^{-1}	Melting point, $^\circ\text{C}$
Pb	48	327
Ag	106	961
Cu	120	1083

Answer: 10^{-4} to 10^{-5} .

- 6.24** Draw a Burgers circuit around a dislocation of symbol \top and determine its Burgers vector.

- 6.25** In a neat sketch, draw Burgers circuits and determine Burgers vectors at two different locations of the same curved dislocation line.

- 6.26** For edge and screw dislocations, list in a tabular form

- (i) the symbols,
- (ii) the angular relationship between the Burgers vector \mathbf{b} and \mathbf{t} vector, and
- (iii) the angular relationship between the direction of motion and the applied stress.

- 6.27** Calculate the dislocation energy per m^3 of FCC copper with a dislocation density of 10^{10} m^{-2} . The shear modulus of copper is 45 GN m^{-2} and the lattice parameter is 3.61 Å .

Answer: 14.7 J m^{-3} .

- 6.28** During a grain growth process, the average grain diameter is initially 0.03 mm . Estimate the maximum possible decrease in the energy during grain growth. The specific grain boundary energy is 0.5 J m^{-2} .

Answer: 50 kJ m^{-3} .

12. The stacking which has a stacking fault in the following is
 A. ...ABCABABCABABCAB...
 B. ...ABCABABCABCABC...
 C. ...ABABABCABABABCABABABC...
 D. ...ABCABABCABABCABABCABABCAB...
13. The stackings with the closest packing of equal-sized spheres are
 A. ...ABBA... B. ...ABCCAB...
 C. ...ACBACB... D. ...ABABCABABC...
14. The local stacking arrangement at a stacking fault in a HCP crystal
 A. ...ABCABABC... B. ...ABABCABABAB...
 C. ...ABABCABABC... D. ...ABCABABCABABC...
15. The tilt angle of a tilt boundary in BCC iron ($a = 2.87 \text{ \AA}$) with edge dislocations 7500 \AA apart is
 A. 0.04° B. 0.2° C. 0.02° D. 0.33 rad
16. The maximum possible decrease in energy during grain growth in Cu (grain boundary energy = 0.5 J m^{-2}) of initial grain diameter of 0.3 mm is
 A. 0.5 kJ m^{-3} B. 2.5 kJ m^{-3} C. 5 kJ m^{-3} D. 10 kJ m^{-3}

Answers

- | | | | | |
|-------|-------|----------|---------|-------|
| 1. B | 2. B | 3. A | 4. A, D | 5. A |
| 6. A | 7. A | 8. A | 9. D | 10. C |
| 11. A | 12. B | 13. C, D | 14. B | 15. C |
| 16. C | | | | |

Source for Experimental Data

Institute of Metals, London, *Dislocations and Properties of Real Materials* (1985).

Suggestions for Further Reading

- J.P. Hirth and J. Lothe, *Theory of Dislocations*, McGraw-Hill, New York (1984).
 A. Kelly, G.W. Groves and P. Kidd, *Crystallography and Crystal Defects*, John Wiley, Chichester, UK (2000).

The study of phase relationships plays an important role in the understanding of the properties of materials. Phase diagrams are maps that give the relationships between phases *in equilibrium* in a system as a function of temperature, pressure and composition. The phases as determined from the microstructure of a material may or may not correspond to the equilibrium phases indicated by the phase diagram. In fact, the thermal treatment given to a material often results in phases other than those in the diagram. This fact, however, does not diminish the importance of the study of phase diagrams as a first step towards gaining an insight into the control of microstructure.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Temperature	kelvin	K	°C, °F
Pressure	pascal	Pa	atm, psi, kbar
	or	or	
	newton per square metre	N m ⁻²	

The compositions given in the text are weight %, unless stated otherwise.

7.1 The Phase Rule

The phase rule enunciated by Gibbs has a simple form:

$$F = C - P + 2 \quad (7.1)$$

where F is the degrees of freedom, C is the number of components and P is the number of phases in equilibrium in a system. Let us examine each of these terms in some detail.

In the classical definition, a *phase* is a physically distinct, chemically homogeneous and mechanically separable region of a system. The various states of aggregation of matter, namely, the gaseous, the liquid and the solid states, form separate phases. The gaseous state is always a single phase, as the atoms (or molecules) in the gas are mixed at the atomic (or molecular) level. A liquid solution is also a single phase. For example, if salt is dissolved in water, the water molecules, the sodium ions and the chlorine ions are mixed at the atomic level in the solution. A liquid mixture (e.g. oil and water), on the other hand, forms two separate phases as there is no mixing at the molecular level. In the solid state, different chemical compositions and different crystal structures are possible so that a solid may consist of several phases. For the same composition, different crystal structures represent different phases. A solid solution has the atoms mixed at the atomic level within the unit cell and is therefore a single phase.

The *components* of a system may be elements, ions or compounds. They refer to the independent chemical species that comprise the system. In the ice–water–steam system, the component is H_2O . In the Cu–Ni system, the elements Cu and Ni are the components, whereas in the Al_2O_3 – Cr_2O_3 system, the two oxides can be taken to be components. In the Fe–C system, iron and graphite can be the components, but it may be often convenient to choose iron and iron carbide (Fe_3C) as the components.

The *variables* of a system include the two external parameters, *temperature* and *pressure*. Within the system, there are variables that specify the compositions of the phases present. Compositions are expressed as weight (or atom %) so that the number of variables required to specify completely the composition of a phase is $(C - 1)$, where C is the number of components in the system. Knowing the percentage of $(C - 1)$ components automatically fixes the percentage of the last component. If there are P phases in a system, the total number of composition variables is $P(C - 1)$. Including the two external variables (pressure and temperature), the total number of variables is $P(C - 1) + 2$. The number of *independent variables* among these gives the degrees of freedom F . Clearly, this number cannot be more than the total number of variables:

$$F = C - P + 2 \leq P(C - 1) + 2 \quad (7.2)$$

It is easy to see from Eq. (7.2) that, when only one phase is present in a system, the degrees of freedom are equal to the total variables. As the number of phases increases, the degrees of freedom decrease. The degree of freedom cannot be less than zero so that we have an upper limit to the number of phases that can exist in equilibrium in a given system.

Example 7.1 What are the degrees of freedom of a system of two components when the number of phases is one, two, three, and so on?

Solution For $C = 2$, we can list the total variables and the degrees of freedom as follows:

No. of phases	Total variables $P(C - 1) + 2$	Degrees of freedom $C - P + 2$
1	3	3
2	4	2
3	5	1
4	6	0

The system cannot have more than four phases in equilibrium.

Phase diagrams are classified on the basis of the number of components in the system. Single-component systems have unary diagrams, two-component systems have binary diagrams, three-component systems give rise to ternary diagrams, and so on.

7.2 Single-component Systems

In single-component systems, there is no composition variable and the only other variables are temperature and pressure. The phase diagram for iron is shown in Fig. 7.1. Pressure is plotted on the x -axis and temperature on the y -axis. The pressure axis starts from 0.1 MPa but can be extended to lower pressures. In phase diagrams, solid phases are denoted by Greek alphabets α , β , γ , etc.

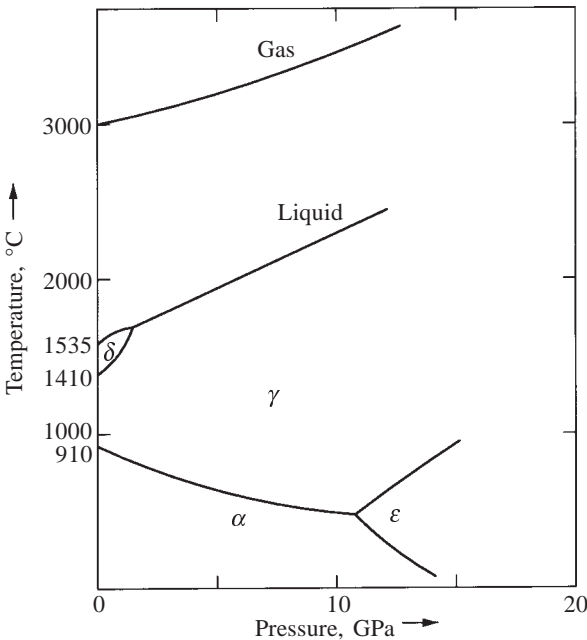


Fig. 7.1 Pressure-temperature diagram for the one-component system of iron.

Consider first the single-phase regions on the diagram such as gas, liquid and several crystal forms of iron. From the phase rule [Eq. (7.1)] for $C = 1$ and $P = 1$, we have $F = 2$. Therefore, in a single-phase region, both temperature and pressure can be varied independently within the limits prescribed by the boundaries of the region.

When two phases are in equilibrium, $F = 1$, that is, either temperature or pressure can be varied independently, *but not both*. Two-phase equilibrium exists *along the phase boundaries* of Fig. 7.1. When we are at a point on the phase boundary if we change the temperature by an arbitrary amount and thereby exercise a degree of freedom in order to preserve the two-phase equilibrium, we need to change the pressure by such an amount that we return to a point on the phase boundary.

Three-phase equilibrium exists at points on the phase diagram, where three phase boundaries meet. Two such points are shown in Fig. 7.1. Such points are called *triple points*. Here, $F = 0$; neither pressure nor temperature can be varied arbitrarily. Three phases will co-exist at only one particular combination of pressure and temperature. If we alter pressure or temperature from the fixed triple-point value, one or two of the phases will disappear.

The equilibrium crystal form of iron at ambient temperature and pressure is BCC (α). On heating to 910°C, α changes over to the FCC (γ) form*. On heating to 1410°C, the γ iron changes over to the BCC (δ) form. As the pressure is increased, the $\alpha \rightarrow \gamma$ transition temperature is lowered, whereas the $\gamma \rightarrow \delta$ transition temperature is increased. Recent high pressure experiments have shown that, when a pressure of about 15 GPa (~150 000 atm) is applied at room temperature, the BCC iron transforms to the HCP (ϵ) phase, as illustrated in Fig. 7.1.

7.3 Binary Phase Diagrams

Two-component systems have binary phase diagrams. Apart from temperature and pressure, we have one composition variable for each of the phases in equilibrium. We then need a three-dimensional diagram to plot the variations in pressure, temperature and composition. In order to simplify the representation of the phase relationships on paper, binary phase diagrams are usually drawn at atmospheric pressure, showing variations in temperature and composition only. Pressure changes often produce no significant effect on the equilibrium and, therefore, it is customary to ignore the pressure variable and the vapour phase. In such cases, one of the variables has been arbitrarily omitted and the phase rule for the condensed phases (solid and liquid phases only) is written in a modified form as

$$F = C - P + 1 \quad (7.3)$$

*Originally, it was believed that a β form of iron also existed, which was later disproved. To avoid confusion, the original naming sequence of the solid phases has been retained with the β phase deleted.

Example 7.2 At atmospheric pressure (pressure arbitrarily chosen), a material of unknown composition shows four phases in equilibrium at 987 K. What is the minimum number of components in the system?

Solution As pressure is arbitrarily chosen, we can use the modified form of the phase rule as given in Eq. (7.3). The minimum number of components corresponds to the minimum degree of freedom, which is zero. Taking $F = 0$, we get

$$0 = C - 4 + 1, \quad C = 3.$$

The simplest binary phase diagram is obtained for a system exhibiting complete liquid solubility as well as solid solubility. The two components dissolve in each other in all proportions both in the liquid and the solid states. Clearly, the two components must have the same crystal structure besides satisfying the other Hume Rothery's conditions for extensive solid solubility, see Sec. 5.3. Cu–Ni, Ag–Au, Ge–Si and Al_2O_3 – Cr_2O_3 are examples of such systems.

Figure 7.2 shows the phase diagram of Al_2O_3 – Cr_2O_3 . Pure Al_2O_3 and pure Cr_2O_3 form the left and the right end of the composition axis. They are arranged in alphabetical order from left to right. The composition is read as per cent of Cr_2O_3 starting from 0% at left and going to 100% at the right end. Alternatively, the composition can be read as per cent of Al_2O_3 from right to left. Temperature is shown along the y-axis.

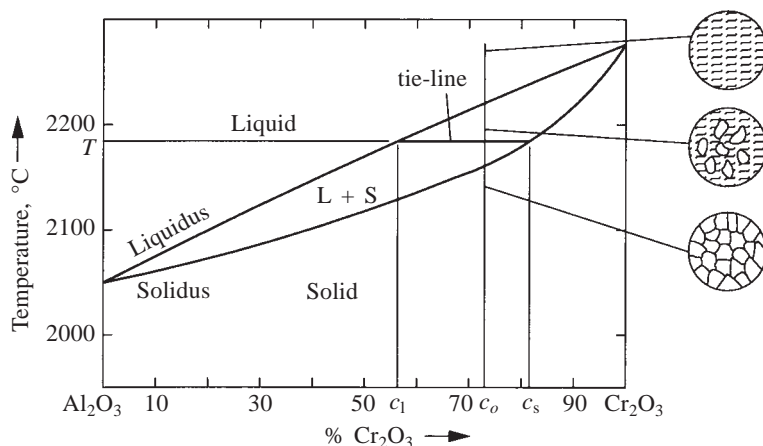


Fig. 7.2 The Al_2O_3 – Cr_2O_3 phase diagram. Microstructural changes in an overall composition c_0 are sketched on the right.

There are only two phases on the phase diagram, the liquid and the solid phases. The single-phase regions are separated by a two-phase region ($L + S$), where both liquid and solid co-exist. In all binary phase diagrams, a two-phase region separates single-phase regions, as given by the “1-2-1” rule. As we move from a single-phase region (1), we cross into a two-phase region (2), and then

again into a single-phase region (1)^{**}. The phase boundary between the liquid and the two-phase region is called the *liquidus*. The boundary between the solid and the two-phase region is called the *solidus*.

When only one phase is present, the composition axis gives the composition of that phase directly. When two phases are present, the compositions of the phases are not the same. They should be read according to the following convention: At the temperature of interest T , a horizontal line called the *tie-line* is drawn as shown in Fig. 7.2. *The points of intersection of the tie-line with the liquidus and the solidus give, respectively, the liquid and the solid compositions, c_l and c_s , which are in equilibrium with each other.* Thus, in Fig. 7.2, at 2180°C, for

an overall composition $c_o = 73\% \text{ Cr}_2\text{O}_3$ (27% Al_2O_3), we have

The liquid composition $c_l = 57\% \text{ Cr}_2\text{O}_3$ (43% Al_2O_3), and

The solid composition $c_s = 82\% \text{ Cr}_2\text{O}_3$ (18% Al_2O_3).

The phase rule can be applied to this phase diagram, using the modified form given in Eq. (7.3). For the single-phase region (liquid or solid), from Eq. (7.3), $F = 2 - 1 + 1 = 2$. So, both temperature and the composition of the phase can be independently varied (within limits). In the two-phase region, $F = 2 - 2 + 1 = 1$.

Here, we have three variables:

- (i) Temperature
- (ii) Composition of the liquid phase
- (iii) Composition of the solid phase

As $F = 1$, only one of these three is independent. If we arbitrarily choose the temperature, the compositions of the two phases are automatically fixed and are given by the ends of the tie-line drawn at that temperature. If we specify the composition of one of the phases arbitrarily, the temperature and the composition of the other phase are automatically fixed. There is no three-phase equilibrium in systems exhibiting complete solid solubility.

Many pairs of elements and compounds are unlikely to satisfy the conditions for complete solid solubility. For instance, the size difference between two atoms or ions can be appreciably more than 15%, as the table of atomic and ionic radii indicates (see back inside cover of the book). Similarly, the other conditions for extensive solubility may not be satisfied. The solid solubility is therefore limited in a number of binary systems. But it is never zero. However unfavourable the conditions for solid solubility are, a very small quantity of any component will always dissolve in another component as this increases the configurational entropy and lowers the free energy of the crystal, recall Problem 2.10. The solubility may be so small that, for all practical purposes, only the pure component may be shown on the phase diagram.

When solid solubility is limited and the melting points of the components are not vastly different, a *eutectic phase diagram* usually results. As an example,

^{**}This is true, except when the phase boundary is a horizontal line, corresponding to an invariant temperature.

the Pb–Sn phase diagram is shown in Fig. 7.3. As there is complete liquid solubility, the liquid phase extends over all compositions above the melting temperatures of the components. The solid phase at the left end is the lead-rich

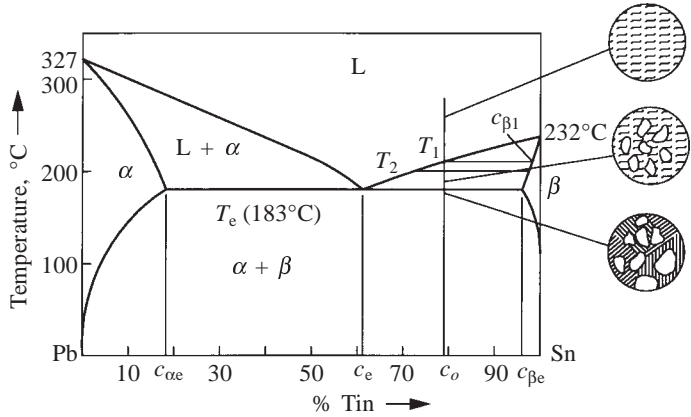


Fig. 7.3 The lead–tin phase diagram. Microstructural changes for composition c_o are sketched.

α which dissolves only a limited amount of tin. This solubility decreases with decreasing temperature. This limit of the solid solubility is indicated by the phase boundary between α and $\alpha + \beta$, called *solvus*. The solid solution phase at the right end is the tin-rich β , with only a very small quantity of lead dissolved in it. The phase boundaries on this diagram are as follows:

- Liquidus I: boundary between L and $L + \alpha$ regions
- Liquidus II: boundary between L and $L + \beta$ regions
- Solidus I: boundary between α and $L + \alpha$ regions
- Solidus II: boundary between β and $L + \beta$ regions
- Solvus I: boundary between α and $\alpha + \beta$ regions
- Solvus II: boundary between β and $\alpha + \beta$ regions

The three two-phase regions are separated by a horizontal line corresponding to the temperature T_e called the *eutectic temperature*. Below the eutectic temperature, the material is fully solid for all compositions. The composition which remains fully liquid up to the eutectic temperature during cooling is called the *eutectic composition* c_e . At the eutectic temperature, the following *eutectic reaction* takes place:



Here, cooling refers to heat being extracted from the system at the eutectic temperature, and heating refers to heat being added to the system. As there is some heat evolution or absorption during the reaction, it is possible to add or subtract heat at constant temperature. The eutectic horizontal is used as a tie-line, the ends of which give the compositions of the α and β phases at this temperature. Summarizing the eutectic characteristics of the Pb–Sn system, we have

eutectic temperature $T_e = 183^\circ\text{C}$,
 composition of liquid $c_e = 62\% \text{ Sn (38\% Pb)}$,
 composition of α , $c_{\alpha e} = 18\% \text{ Sn (82\% Pb)}$, and
 composition of β , $c_{\beta e} = 97\% \text{ Sn (3\% Pb)}$.

The phase rule is readily applied in the single-phase and the two-phase regions of the diagram, on the same lines as discussed for the $\text{Al}_2\text{O}_3\text{--Cr}_2\text{O}_3$ system. At the eutectic temperature T_e , three phases are in equilibrium. From Eq. (7.3), $F = 2 - 3 + 1 = 0$. The eutectic temperature T_e and the compositions of the three phases, c_e , $c_{\alpha e}$ and $c_{\beta e}$ are all fixed and none of them can be varied arbitrarily. On slightly increasing the temperature above T_e , either one or both of α and β phases would disappear. On slight decrease of temperature below T_e , the liquid phase would transform as per reaction (7.4) to a mixture of α and β . To denote the zero degree of freedom, the eutectic reaction is called an *invariant reaction*. The eutectic temperature is known as an *invariant temperature*.

A similar invariant reaction occurring entirely in the solid state, where the liquid phase is replaced by a third solid phase γ , is called a *eutectoid reaction*:



The corresponding invariant temperature is called the *eutectoid temperature*.

When the melting points of the components are vastly different from each other, a *peritectic phase diagram* may be formed. As an example, the Ag–Pt phase diagram is shown in Fig. 7.4. The melting points of the components differ

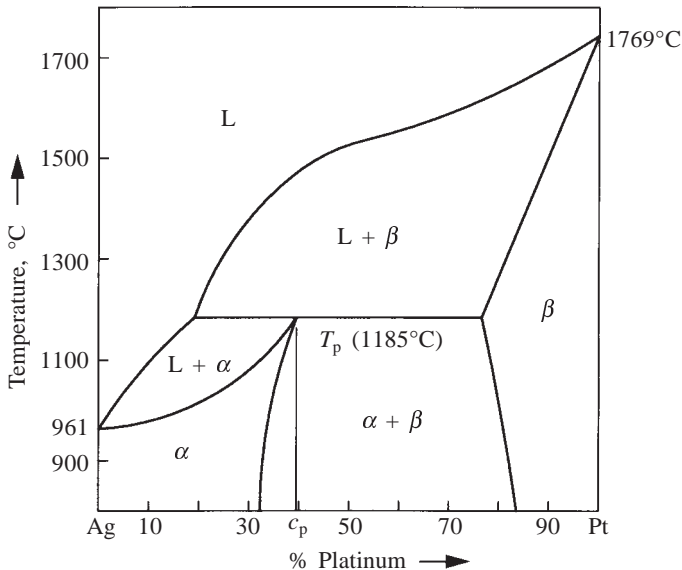


Fig. 7.4 The silver–platinum phase diagram.

by more than 800°C. At the peritectic temperature T_p , the following invariant reaction occurs:



In contrast to the eutectic reaction, the peritectic reaction is between a liquid and a solid phase to yield on cooling a single solid phase. The peritectic composition c_p is the composition of the product solid phase. The peritectic horizontal is the tie-line which defines the composition of the liquid phase and the β phase at the peritectic temperature.

Analogous to the eutectoid reaction, we have a peritectoid reaction, where two solids react on cooling to produce a third solid phase:



The invariant reactions that we have considered up till now are summarized in Table 7.1.

TABLE 7.1
Invariant Reactions

Name of the reaction	Details of the reaction	Phase boundaries at the invariant line
Eutectic	cooling \rightarrow $L \leftrightarrow \alpha + \beta$ \leftarrow heating	
Eutectoid	cooling \rightarrow $\gamma \leftrightarrow \alpha + \beta$ \leftarrow heating	
Peritectic	cooling \rightarrow $L + \beta \leftrightarrow \alpha$ \leftarrow heating	
Peritectoid	cooling \rightarrow $\gamma + \beta \leftrightarrow \alpha$ \leftarrow heating	

Binary diagrams have other invariant reactions, such as monotectic and syntectic reactions. In addition, many binary systems have compounds at certain ratios of the components, e.g., Mg_2Pb in the Mg–Pb system. Some of these compounds exist over a range of composition rather than at a fixed value. Further details about binary phase diagrams and systems with more than two components can be found in the references listed at the end of this chapter.

7.3 Microstructural Changes during Cooling

The microstructural changes that occur on cooling an overall composition c_o in the $\text{Al}_2\text{O}_3\text{--Cr}_2\text{O}_3$ diagram are illustrated in Fig. 7.2. Above the liquidus, the material is a single phase (liquid). On cooling, we cross the liquidus and enter the two-phase region. At temperature T , solid crystals of composition c_s co-exist with liquid of composition c_l . On further cooling, we cross the solidus and the material is now fully crystallized into a polycrystalline solid of composition c_o .

In the eutectic system of Fig. 7.3, let us consider the cooling of an overall composition c_o . At temperature T_1 , solidification starts with a very small quantity of β crystals of composition $c_{\beta 1}$ coming out of the liquid. At a lower temperature T_2 , the amount of β phase has increased. Its composition is now different and is given by the tie-line drawn at this temperature. Just above the eutectic temperature, the quantity of β crystals has further increased, their composition now being $c_{\beta e}$. The microstructure of the alloy at this stage would be composed of β crystals floating in the liquid of eutectic composition, as shown in Fig. 7.3. On cooling just below the eutectic temperature, the liquid transforms according to the eutectic reaction (7.4) to yield a fine mixture of thin plate-like crystals of α and β , see Fig. 7.3. The microstructure below the eutectic temperature is re-sketched in Fig. 7.5. The large equiaxed crystals of *proeutectic* β , that is, β

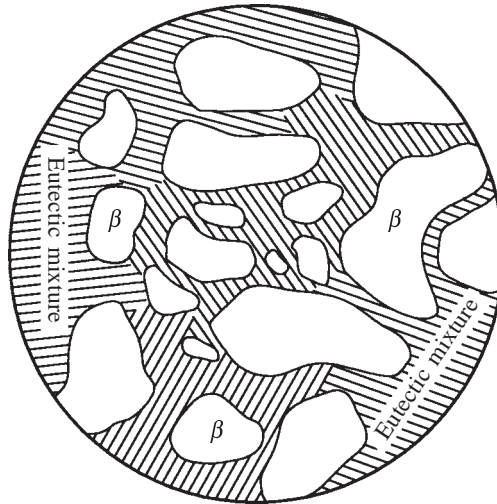


Fig. 7.5 Microstructure of a hypereutectic Pb–Sn alloy, showing the proeutectic β crystals in a matrix of the eutectic mixture (schematic).

formed before the eutectic reaction took place, are clearly seen. The matrix in which they are embedded is the eutectic mixture consisting of thin, parallel plates of alternate α and β crystals. Since the eutectic region is clearly distinguishable as a separate entity under the microscope, it is called a *micro-constituent*. It should not, however, be called the eutectic phase as it is not a single phase.

If the overall composition of the alloy is c_e , the entire solidification takes place at the eutectic temperature and the microstructure in this case has only the eutectic mixture, as illustrated in Fig. 7.6. If the overall composition is to the left

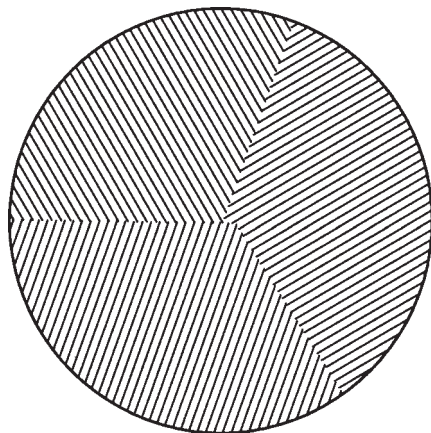


Fig. 7.6 Microstructure of a Pb-Sn alloy of eutectic composition (schematic).

of the eutectic composition, then the microstructure consists of large equiaxed crystals of proeutectic α that form prior to the eutectic reaction. These α crystals get embedded in the eutectic matrix that forms at the eutectic temperature. Alloy compositions to the left of the eutectic mixture are called *hypoeutectic alloys*, and those to the right are called *hypereutectic alloys*.

What we have described above corresponds to an extremely slow cooling so that the concerned phases have the equilibrium composition as given by the tie-line at every temperature. This means that the composition of the equiaxed β crystals of the hypereutectic alloy of Fig. 7.3 should change continuously from $c_{\beta 1}$ to $c_{\beta e}$, as we cool from T_1 to T_e . Atoms must move across the liquid-crystal boundary (from or to the liquid) to bring about these compositional adjustments. Atom movements in the solid are so slow that these adjustments are rarely complete. At ordinary cooling rates, the crystals show a gradation of composition from $c_{\beta 1}$ at the centre of the crystals, where solidification first started, to $c_{\beta e}$ at the peripheral rims of the crystals which solidified just above the eutectic temperature. This compositional inhomogeneity is called *coring*. Coring can be identified by the contrast a cored crystal shows on chemical etching.

Sometimes crystals grow *preferentially* along certain crystallographic directions. Frequently, in cubic crystals, the three mutually perpendicular $\langle 100 \rangle$ directions are the preferred directions of growth. During the early stages of solidification, the shape of the crystal is like a fir tree, with arms protruding out in three directions, as seen in Fig. 7.7. This structure is called the *dendritic structure*. The liquid between the dendritic arms undergoes solidification subsequently. The compositional differences caused by the coring effect enable the identification of the dendrites even after complete solidification.

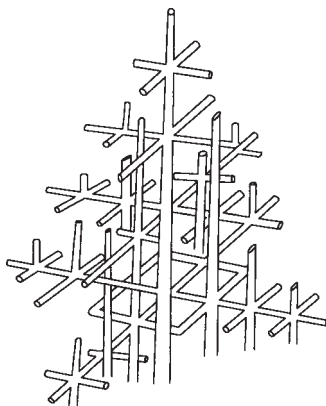


Fig. 7.7 The dendritic structure.

7.5 The Lever Rule

As already described, the compositions of the two co-existing phases of a binary system are given by the tie-line rule. For all overall compositions that lie on the tie-line, the compositions of the two phases remain the same, c_s and c_l in Fig. 7.2. A little reflection will show that this will be possible only if the relative amounts of the co-existing phases change, as the overall composition is varied along the tie-line. *Lever rule* derived from mass balance gives the relative amounts of the co-existing phases. It is applied as follows: *The tie-line is treated as a lever arm, with the fulcrum at the overall composition.* For the arm to be horizontal, the weight to be hung at each end must be proportional to the arm length on the other side of the fulcrum. The “weight” at each end corresponds to the amount of the phase at that end. At temperature T and overall composition c_o in Fig. 7.2, the relative amounts of the liquid and the solid phases are determined as follows. Figure 7.8 illustrates the procedure. Expressing the

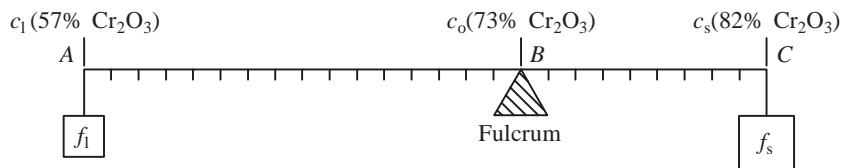


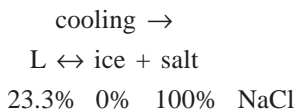
Fig. 7.8 Illustration of the lever rule.

weight fractions of liquid and solid as f_l and f_s , we obtain

$$f_l = \frac{CB}{CA} = \frac{c_s - c_o}{c_s - c_l} = \frac{82 - 73}{82 - 57} = 0.36 \quad (7.8)$$

$$f_s = \frac{BA}{CA} = \frac{c_o - c_l}{c_s - c_l} = \frac{73 - 57}{82 - 57} = 0.64 \quad (7.9)$$

Example 7.3 The H_2O – NaCl system has the following eutectic reaction occurring at -21°C :



How much pure water can be extracted from sea water (3.5% NaCl) by cooling to -20.9°C ?

Solution -20.9°C is just above the eutectic temperature of -21°C . Here, liquid of eutectic composition (23.3% NaCl) is in equilibrium with pure ice (0% NaCl). The ice crystals can be separated from the liquid and melted into water. With the fulcrum at the overall composition of sea water (3.5% NaCl) and the lever arm extending to 23.3% NaCl on one side to 0% on the other side, we obtain

$$f_{\text{ice}} = (23.3 - 3.5)/(23.3 - 0) = 19.8/23.3 = 0.85$$

In the above examples, only two phases are present. The lever rule cannot be applied at the eutectic or the peritectic temperature, where there are three phases in equilibrium and an isothermal reaction, changing the relative amounts of the phases, can occur. It can be applied just above or just below the invariant temperature. By using the lever rule, it is possible to estimate

- (i) the fraction of a proeutectic phase,
- (ii) the fraction of the eutectic mixture, and
- (iii) the fraction of a phase that forms part of the eutectic mixture.

Referring to Fig. 7.3, just below the eutectic temperature, for the overall composition c_o , the fraction of proeutectic β , $f_{\text{pro } \beta}$, is determined as follows: The fulcrum of the lever is positioned at the overall composition. One end of the lever arm ends at the phase boundary corresponding to β , that is, at the composition $c_{\beta e}$. The other end of the lever arm extends up to the average composition of the eutectic mixture c_e . We can then write

$$f_{\text{pro } \beta} = \frac{c_o - c_e}{c_{\beta e} - c_e} \quad (7.10)$$

The fraction of the eutectic mixture is given by

$$f_{\text{eut}} = \frac{c_{\beta e} - c_o}{c_{\beta e} - c_e} \quad (7.11)$$

In order to determine the total β , which is the sum of the proeutectic β and the β in the eutectic mixture, the fulcrum as before remains at the overall composition. The ends of the lever arm, however, extend up to $c_{\beta e}$ on the β side and up to $c_{\alpha e}$ on the α side.

$$f_{\text{total } \beta} = \frac{c_o - c_{\alpha e}}{c_{\beta e} - c_{\alpha e}} \quad (7.12)$$

Example 7.4 What is the fraction of β that forms part of the eutectic mixture in the alloy of overall composition of c_o in Fig. 7.3?

Solution The fraction of the β phase that forms part of the eutectic mixture is equal to the total β in the material minus the proeutectic β . From Eqs. (7.10) and (7.12),

$$\begin{aligned} f_{\text{eut } \beta} &= f_{\text{total } \beta} - f_{\text{pro } \beta} \\ &= \frac{c_o - c_{\alpha e}}{c_{\beta e} - c_{\alpha e}} - \frac{c_o - c_e}{c_{\beta e} - c_e} \end{aligned}$$

This fraction can also be expressed as the product of two fractions as follows:

$$\begin{aligned} f_{\text{eut } \beta} &= f_{\text{eut in this alloy}} \times f_{\beta \text{ in eut alloy}} \\ &= \frac{c_{\beta e} - c_o}{c_{\beta e} - c_e} \times \frac{c_e - c_{\alpha e}}{c_{\beta e} - c_{\alpha e}} \end{aligned}$$

It is left as an exercise for the student to show that both these expressions for fraction of eutectic β are equal.

7.6 Summary of Phase Diagram Rules

As a review, the rules pertaining to phase diagrams are summarized as follows:

7.6.1 The Phase Rule

The Gibbs phase rule states that

$$F = C - P + 2$$

where F is the degrees of freedom, P the number of phases, and C the number of components. A phase is a region of a system with the same structure and uniform composition and differs from other regions of the system either in structure and/or composition. The component refers to the individual chemical species. The degrees of freedom are the number of independent variables associated with the equilibrium. It is equal to the total variables, when there is only one phase in the system. It decreases as the number of phases increases. The maximum number of phases in equilibrium in a system corresponds to $F = 0$.

7.6.2 The Tie-Line Rule

The tie-line rule is applied to determine the compositions of two co-existing phases in a binary phase diagram. The tie-line is a horizontal line drawn at the temperature of interest within the two-phase region. If, for example, a liquid phase and a solid phase co-exist at a temperature T , the intersection of the

tie-line drawn at that temperature with the liquidus gives the composition of the liquid, and the intersection with the solidus gives the composition of the solid. The tie-line rule is not concerned with the fraction of the co-existing phases. It can be applied only in the two-phase region. Note that, for all overall compositions that lie on the tie-line, the compositions of the two co-existing phases remain the same. There is change only in their relative amounts, which is determined by the lever rule.

7.6.3 The Lever Rule

The lever rule gives the fractions of two co-existing phases. The tie-line at the temperature of interest is treated as a lever arm, with the fulcrum at the overall composition. For the lever arm to be horizontal, the weight hung at each end must be proportional to the arm length on the other side of the fulcrum. The “weight” at each end corresponds to the amount of the phase at that end.

It should be noted that the tie-line rule gives the composition of the co-existing phases, whereas the lever rule gives the fractions (or amounts) of the phases. In a binary system of A and B, if a liquid of 35% A (65% B) is co-existing with a solid of 75% A (25% B), for an overall composition of 40% A, the fraction of the liquid is given by

$$f_l = \frac{75 - 40}{75 - 35} = 0.875$$

The fraction of the solid $f_s = 1 - f_l = 1 - 0.875 = 0.125$.

Stating the tie-line rule and the lever rule together, for the above example, we get

$$\begin{aligned} \text{fraction of liquid with 35\% A} &= 0.875 \\ \text{fraction of solid with 75\% A} &= 0.125 \\ 0.875 \times 35 + 0.125 \times 75 &= 40\% \text{ A} \\ &= \text{overall composition of the alloy} \end{aligned}$$

The lever rule cannot be applied at an invariant temperature, where three phases are in equilibrium. It can, however, be applied just above or below the invariant line. The lever rule can be used to calculate the fraction of a proeutectic phase, the fraction of a eutectic mixture or the fraction of a phase that forms part of the eutectic mixture. In such calculations, the fulcrum is always at the overall composition. If the fraction of a eutectic mixture is to be calculated, one end of the lever arm must end at the average composition of the eutectic mixture, that is, at the eutectic composition.

7.6.4 The 1-2-1 Rule

The 1-2-1 rule is an aid to label the phase fields in a binary phase diagram. As we move from a single-phase region (1), we cross into a two-phase region (2), and then again into a single-phase region (1). The 1-2-1 rule is not applicable across an invariant horizontal line.

Note that the boundaries of a two-phase region always meet at a point on reaching a pure component. This is consistent with the fact that a pure component undergoes a phase change such as crystallization at a constant temperature.

7.7 Some Typical Phase Diagrams

Phase diagrams may look somewhat complicated if there is more than one invariant reaction in the system. As examples, we consider a few more typical phase diagrams.

7.7.1 Magnesita–Alumina System

The $\text{MgO}-\text{Al}_2\text{O}_3$ phase diagram is shown in Fig. 7.9. It has two eutectic reactions, and an intermediate phase, $\text{MgO} \cdot \text{Al}_2\text{O}_3$. There is very little terminal

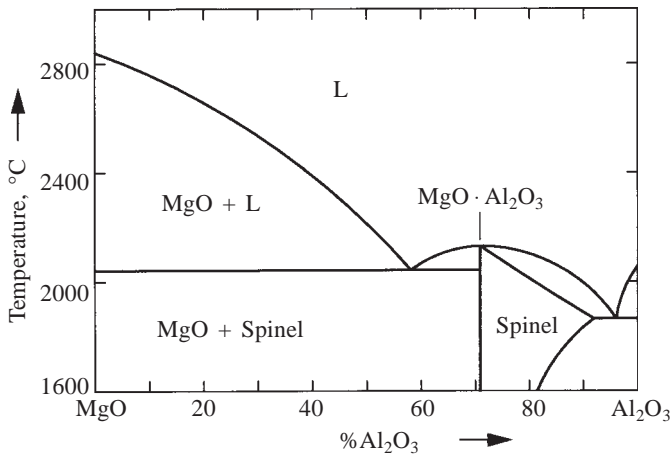


Fig. 7.9 The $\text{MgO}-\text{Al}_2\text{O}_3$ phase diagram.

solubilities, that is, Al_2O_3 dissolves only a negligible amount of MgO and vice versa. So, for all practical purposes, the pure end phases are the same as components. This is in contrast to the $\text{Pb}-\text{Sn}$ system, where we have Pb and Sn as the two components and α and β as the two phases near the pure ends, see Fig. 7.3. The $\text{MgO}-\text{Al}_2\text{O}_3$ system can also be considered as two separate phase diagrams, the components for one being MgO and $\text{MgO} \cdot \text{Al}_2\text{O}_3$, and for the other being $\text{MgO} \cdot \text{Al}_2\text{O}_3$ and Al_2O_3 .

Note that a small quantity of MgO in Al_2O_3 lowers its melting point appreciably, even though MgO has a very high melting point. Such a phenomenon can cause problems when oxides are used as high temperature refractories. High quality pure refractories with a high softening temperature can be ruined due to the presence of impurities that lower their melting point.

7.7.2 Copper–Zinc System

Commercially important Cu–Zn alloys, *brasses*, have a composition up to about 40% Zn. The Cu–Zn phase diagram from 0 to 60% Zn is shown in Fig. 7.10a. Zinc melts at 419°C and copper at 1083°C. This large difference in the melting points results in a number of peritectic reactions in the system, two of which fall in the range shown in Fig. 7.10a. The phases present in brass at room temperature as a function of zinc content are listed as follows:

α	$\alpha + \beta$	β
0–35% Zn	35–46.6% Zn	46.6–50.6% Zn

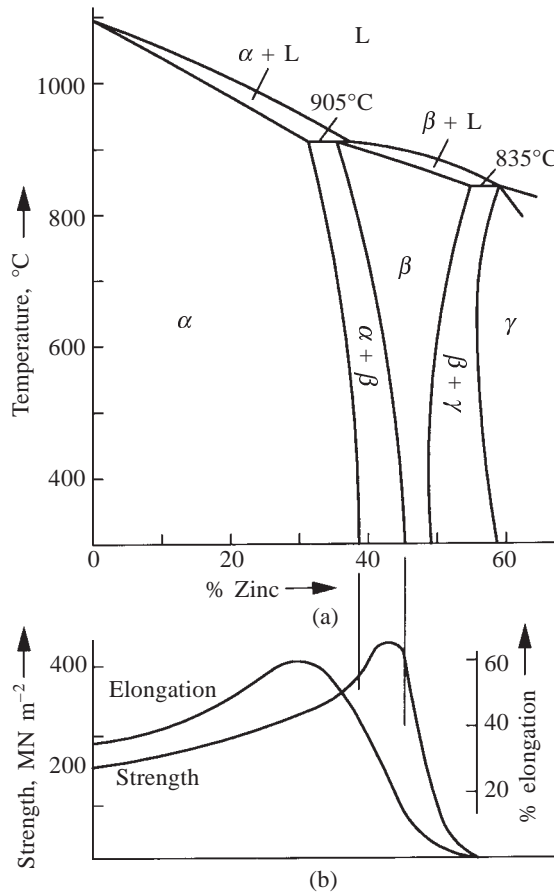


Fig. 7.10 (a) The copper–zinc phase diagram up to 60% zinc; (b) the variation in strength and % elongation with zinc content.

Zinc is cheaper than copper and, therefore, brasses are more economical to use than pure copper. In addition, the mechanical properties of copper such as tensile strength and % elongation improve with the addition of zinc, see

Fig. 7.10b. Brasses are widely used in a variety of applications. Where electrical conductivity is an important consideration, pure copper is used.

7.7.3 Iron–Iron-carbide System

Fe and Fe_3C (iron carbide) are the components in the Fe– Fe_3C phase diagram. It is also possible to have a phase diagram with Fe and C (graphite) as the components. Graphite is more stable than Fe_3C . So, the Fe– Fe_3C diagram can be considered to be a metastable phase diagram. Most steels contain only iron carbide and not graphite.

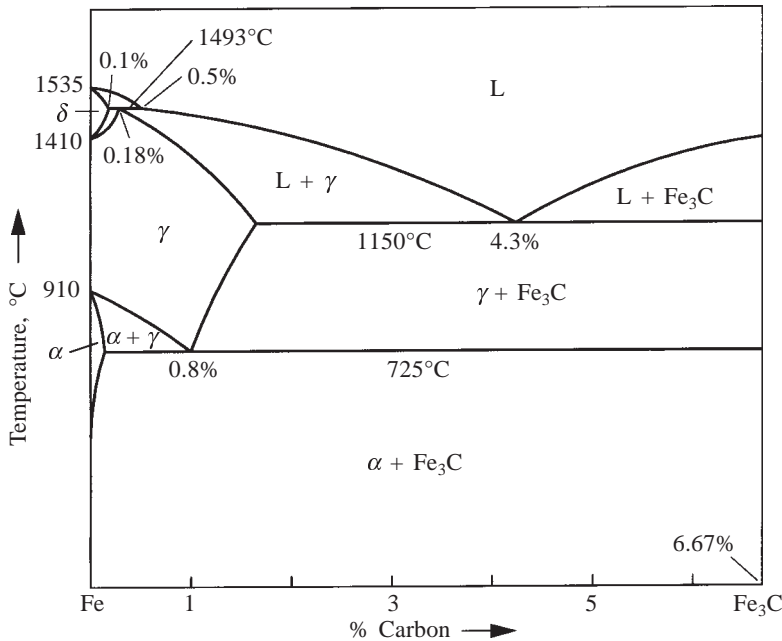


Fig. 7.11 The Iron–Iron-Carbide phase diagram.

In Fig. 7.11, the composition is plotted on the x -axis as weight per cent of carbon. Pure iron melts at 1535°C . A peritectic invariant reaction occurs at 1493°C with the peritectic composition at 0.18% carbon:



The product phase γ is called *austenite*. A eutectic reaction occurs at 1150°C with the eutectic composition at 4.3% carbon:



A eutectoid reaction occurs at 725°C with the eutectoid composition at 0.8% carbon:



Here, austenite (γ) decomposes into two phases—ferrite (α) and cementite (Fe_3C).

Fe–C alloys containing 0–1.4% carbon are called *steels*. The approximate range of carbon content of mild steels is 0–0.3%, for medium carbon steels it is 0.3–0.6%, and for high carbon steels, it is 0.8–1.4%. Here, we will consider the microstructures of slowly cooled steels of different carbon content. When the carbon content is negligible, the structure of steel is essentially polycrystalline α (ferrite), which is the interstitial solid solution of carbon in BCC iron.

When the steel has the eutectoid composition of 0.8% carbon, the microstructure consists of alternate layers of thin, parallel plates of α (ferrite) and Fe_3C (cementite). This eutectoidal mixture is called *pearlite*. Pearlite is a microconstituent, and not a single phase. It is so named as it has an iridescent appearance under the microscope, resembling the mother of pearl, see Fig. 7.12.

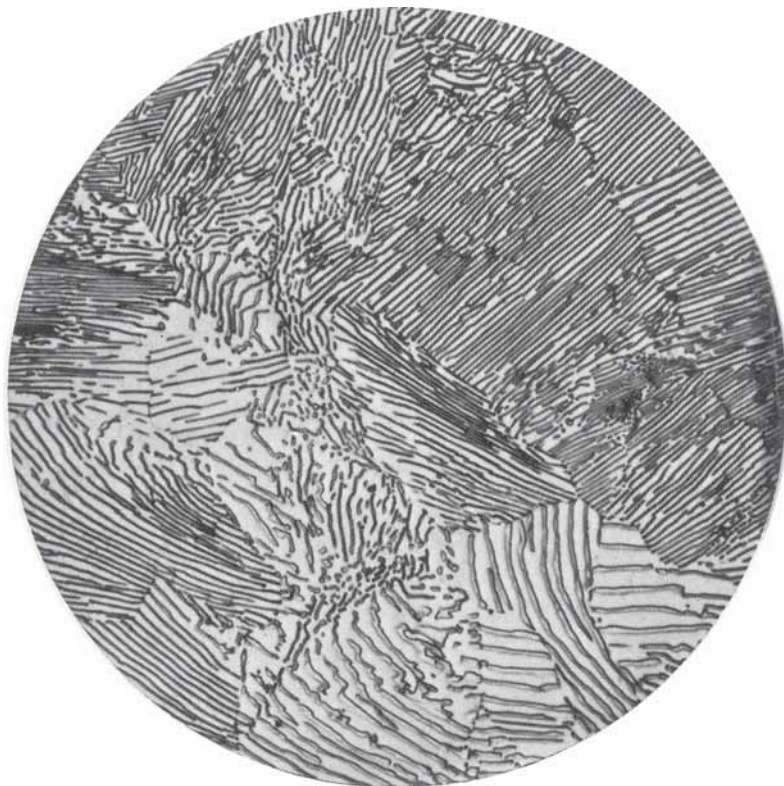


Fig. 7.12 Microstructure of eutectoid steel showing pearlite, magnified 1000 times. (E.C. Bain and H.W. Paxton, *Alloying Elements in Steel*, with permission from ASM International, Materials Park, Ohio.)

Applying the lever rule, with fulcrum at 0.8% carbon and the lever arm extending up to ferrite ($\sim 0.0\%$ carbon) at one end and up to cementite (6.67% carbon) at the other end, we have, in pearlite, the relations

$$f_{\alpha} = (6.67 - 0.8)/(6.67 - 0.0) = 0.88$$

$$f_{\text{Fe}_3\text{C}} = (0.8 - 0.0)/(6.67 - 0.0) = 0.12$$

The approximately-parallel lines in the microstructure indicate the layer arrangement of ferrite and cementite. As the fraction of cementite is less than that of ferrite, the cementite plates are thinner than ferrite in Fig. 7.12. Both the boundaries of a cementite plate are clearly visible in some regions of the microstructure, while in other parts, the cementite plates appear to be just a line, as the two boundaries of a plate are too close to be resolved as separate lines. At lower magnifications or when the pearlite mixture is very fine, even ferrite and cementite may not be resolved as separate platelets and pearlite may appear as just dark regions under the microscope.

Mild steel with 0.2% carbon consists of about 75% of proeutectoid ferrite that forms above the eutectoid temperature and about 25% of pearlite. When the carbon content in the steel is increased, the amount of pearlite increases, until we get the fully pearlitic structure at 0.8% carbon. Beyond 0.8%, high carbon steels contain proeutectoid cementite in addition to pearlite.

Example 7.5 How much proeutectoid ferrite is there in a slowly cooled 0.6% steel? How much eutectoid ferrite is there in the same steel?

Solution Applying lever rule between 0.0 and 0.8% carbon with fulcrum at 0.6% carbon,

$$f_{\text{pro } \alpha} = (0.8 - 0.6)/(0.8 - 0.0) = 0.25$$

$$f_{\text{pearlite}} = 1.0 - 0.25 = 0.75$$

The fraction of ferrite in a eutectoid steel = 0.88

$$f_{\text{eut } \alpha \text{ in this steel}} = 0.75 \times 0.88 = 0.66.$$

In slowly cooled carbon steels, the overall hardness and ductility of the steel are determined by the relative proportions of the soft, ductile ferrite and the hard, brittle cementite. The cementite content increases with increasing carbon content, resulting in an increase of hardness and a decrease of ductility, as we go from low carbon to high carbon steels. Table 7.2 lists typical applications of steels according to carbon content. A rivet (at the top of the list) should have good deformability, and hence has a low carbon content. In contrast, a file (at the bottom of the list) should have high hardness and wear resistance and hence has a high carbon content. Even though we may increase the hardness by increasing the carbon content further, the alloy becomes too brittle to be useful above 1.4% carbon. A rail (in the middle of the table) has 0.6% carbon. It combines some toughness with some hardness and wear resistance.

TABLE 7.2
Some Applications of Plain Carbon Steels

Type of steels	Carbon percentage	Uses
Mild steels	0.0–0.1	Chain links, nails, rivets
	0.1–0.2	Ship hulls, car bodies
	0.2–0.3	Bridges
Medium carbon steels	0.4–0.5	Train axles and wheels
	0.5–0.6	Rails
	0.6–0.7	Chisels
	0.8–0.9	Saw cutters, hammers
High carbon steels*	1.0–1.1	Axes
	1.1–1.2	Razor blades
	1.3–1.4	Scissors, knives, files

*Many of these are used after an appropriate thermal treatment.

Fe–C alloys with more than 2% carbon are called *cast irons*. Consider, for example, the cooling of a cast iron with 3% carbon, see Fig. 7.11. On crossing the liquidus into the (L + γ) region, proeutectic γ (austenite) crystallizes first. On passing through the eutectic temperature, liquid of eutectic composition decomposes to a mixture of austenite and cementite. On further cooling through the eutectoid temperature, the austenite decomposes to pearlite, yielding *white cast iron*. The microstructure of white cast iron consists of cementite and pearlite, see Fig. 7.13. Due to the presence of the large fraction of cementite, white cast iron is very hard and brittle.

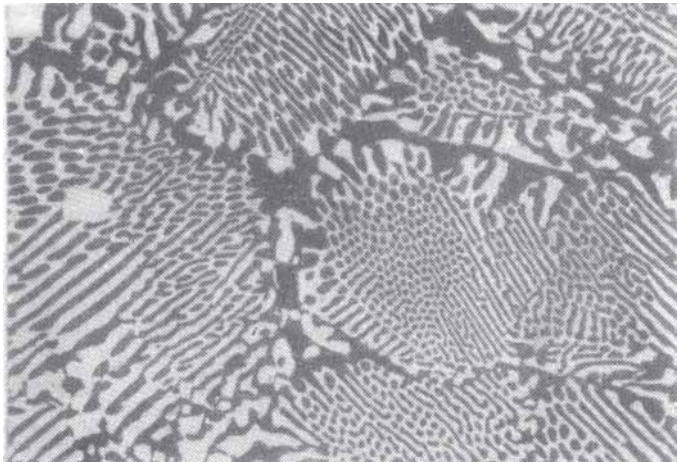


Fig. 7.13 Microstructure of white cast iron, magnified 650 times. White regions are cementite and dark regions are unresolved pearlite. (Y. Lakhtin, *Engineering Physical Metallurgy*, with permission from Mir Publishers, Moscow.)

Depending on the cooling rate and the other alloying elements present in cast iron, the carbon may be present as graphite or cementite. *Gray cast iron* contains graphite in the form of flakes. Slow cooling rates and the presence of silicon promote the formation of graphite. The microstructure of gray cast iron is shown in Fig. 7.14 and consists of graphite flakes in a matrix of ferrite. The



Fig. 7.14 Microstructure of gray cast iron, magnified 140 times. Graphite flakes are embedded in a matrix of ferrite. (A.L. Ruoff, *Introduction to Materials Science*, with permission from Prentice Hall, Inc., Englewood Cliffs, New Jersey.)

graphite flakes are sharp at their tips and act as stress raisers. Due to this, gray cast iron is brittle under tensile loads, in spite of the softness of graphite as compared to the very hard cementite present in white cast iron. The brittleness can be avoided by producing the graphite in the form of spherical nodules (which do not have stress-raising sharp ends), as is done in *malleable cast iron* and SG (*spheroidal graphite*) iron.

Malleable cast iron is produced by heat treating white cast iron for prolonged periods at about 900°C and then cooling it very slowly. The silicon content in the alloy must be 1% or less to ensure that cementite and not flaky graphite forms during solidification. The cementite decomposes to the more stable graphite during the subsequent heat treatment. Graphite appears in the heat-treated microstructure as approximately spherical particles of *temper carbon*. SG iron (also known as *nodular iron*) is produced by making certain alloy additions such as Mg or Ce to molten iron. Here, the silicon content must be about 2.5% to promote graphitization. The alloy additions have the effect of modifying the growth rate of graphite from the melt to be more or less equal in all directions, so that nodules (and not flakes) of graphite are produced. No subsequent heat treatment is required for SG iron.

7.7 Other Applications of Phase Diagrams

One of the important uses of a phase diagram is in the *zone refining* of materials. Consider the schematic phase diagram shown in Fig. 7.15. We need to purify (or refine) A, which has B as the impurity component. Let the starting

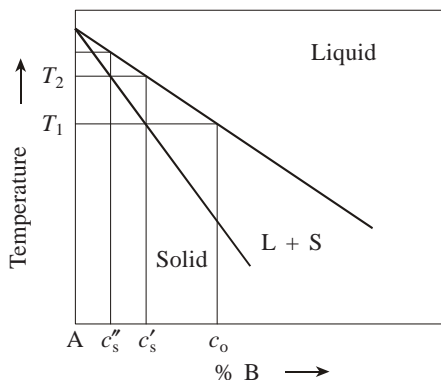


Fig. 7.15 The zone refining process is based on the fact that the solid to crystallize first from a melt is usually purer in A than the liquid.

composition be c_0 . This composition starts to solidify at temperature T_1 . Just below this temperature, a small quantity of solid of composition c'_s , purer in A than c_0 , separates from the liquid. Imagine that we stop cooling further, throw away the liquid part and remelt the left-over solid. This small quantity of liquid is now purer, as its composition is c'_s . This overall composition will now start to solidify at a higher temperature T_2 . Again, the first solid to separate will have a composition c''_s purer in A than c'_s . By repeating this sequence of operations a few times, we can get very pure A, even though the quantity of the purified material will be extremely small compared to the starting material.

In zone refining, this principle of phase separation is used. The material to be purified is in the form of a long rod. At any time, only a small length of this rod is melted with the aid of an induction coil or an electron beam. The coil or beam is moved slowly from one end to the other end of the rod, continuously solidifying the molten zone and remelting fresh material ahead. Surface tension forces are usually strong enough to hold the molten zone in place without the need for a container, which may contaminate the melt. If the zone is passed across many times, each time in the same direction, the material at the starting end becomes much purer than the rest of the rod. In a typical case, ten passes of the molten zone can reduce the impurity level to as low as 10^{-6} times the initial value. This technique to produce ultra high purity materials has enabled us to achieve the sophistication required in the control of impurity concentration in semiconductor crystals.

The depression of the melting point of a pure solid, when another component goes into solution, has its advantages and disadvantages. The problem of softening and melting of refractories due to impurities was referred

to earlier. Some of the advantages are described below. The eutectic reaction in the iron-carbon alloys facilitates the melting of cast irons at a lower temperature than is required for steels. So, the making of cast iron is easier and more economical. In cold countries, the depression of the freezing point of water by the addition of salt is used to melt the ice on roads. Pb–Sn eutectic alloys are useful as soldering material. Low melting eutectic alloys are used as safety devices in fire fighting equipment and petroleum storage tanks. Gold–silicon eutectic finds uses in the manufacture of semiconductor devices. Closing this section of applications in a lighter vein, a teaspoon made out of a four-component eutectic of Pb, Sn, Bi and Cd melts at 70°C, inside a hot cup of tea!

SUMMARY

1. The phase rule gives the degrees of freedom that are permissible for a system of a given number of components and phases in equilibrium.
2. The pressure-temperature diagrams of one component systems have phase boundaries corresponding to two-phase equilibrium and triple points corresponding to three phase equilibrium.
3. Binary phase diagrams are temperature-composition plots, where the pressure variable is omitted.
4. Eutectic and peritectic systems have a fixed (invariant) temperature for the existence of three-phase equilibrium. Here, the compositions of all the three phases are fixed.
5. The solid state analog of the eutectic and the peritectic reactions are the eutectoid and the peritectoid reactions.
6. Lever rule is a simple rule of mass balance, giving the relative amounts of two phases (or microconstituents) for a given overall composition.
7. Brasses are more economical and have better mechanical properties than copper.
8. The microstructures of slowly cooled steels and cast irons correspond to the phases shown in the phase diagram.
9. The zone refining process is based on the fact that the first solid to crystallize in a two-component system is generally purer than the liquid.

PROBLEMS

- 7.1** Water vapour, ice and water are in equilibrium at 0.01°C and at a pressure of 613 Pa. Which of these phases would disappear, if
- (i) the temperature is decreased, and
 - (ii) the pressure is increased?

- 7.2** In a binary phase diagram (pressure omitted), what is the maximum number of phases that can co-exist for at least one degree of freedom?
- 7.3** Give the number of variables and the degrees of freedom at the peritectic temperature of a binary phase diagram.
- 7.4** Trace the microstructural changes on slow cooling from the liquid state of an alloy of any overall composition that undergoes the peritectic reaction in the Ag–Pt system (Fig. 7.4).
- 7.5** Dendritic growth cannot be identified by an etching technique in pure materials. Why?
- 7.6** Show that, for correct mass balance, the relative amounts of two co-existing phases or microconstituents must be as given by the lever rule.
- 7.7** The following data apply to the binary system of A and B:

Melting point of pure A = 1050°C

Melting point of pure B = 1900°C

At 1250°C, the α solid solution (containing 50% of B), the β solid solution (containing 80% of B) and liquid (containing 30% of B) are in three-phase equilibrium. At room temperature, the maximum solubility of B in the α phase is 30% and the maximum solubility of A in the β phase is 10%.

- (i) Sketch the phase diagram of A and B, showing the phases present in each area of the diagram.
- (ii) What is the reaction that takes place at 1250°C?
- (iii) What are the weight fractions of the phases present in a material of overall composition 75% B, at 1251°C, 1249°C and at room temperature (RT)?

Answer: (iii) At 1251°C, $f_l = 0.1$, $f_\beta = 0.9$;
 at 1249°C, $f_\alpha = 0.167$, $f_\beta = 0.833$; and
 at RT, $f_\alpha = 0.25$, $f_\beta = 0.75$.

- 7.8** In the Pb–Sn system, determine the fraction of β phase in an alloy of 80% Sn at 184°C and 182°C.

Answer: 0.51 and 0.78.

- 7.9** In the microstructure of a Pb–Sn solder alloy, it was found that 88% of the area is occupied by the eutectic constituent and the remaining area by the proeutectic β phase. Assuming that the area fractions are equal to the volume fractions, determine the approximate composition of the alloy. The densities of α and β phases are 10 300 and 7300 kg m⁻³.

Answer: 65.7%.

- 7.10** The phase diagram of a binary system of A and B has a three-phase equilibrium at 250°C , with the compositions of α , liquid and β phases equal to 10%, 55% and 95% B. Just below 250°C , find the compositions at which the proeutectic phase is $1\frac{1}{2}$ times the eutectic mixture.

Answer: 28%, 79%.

- 7.11** Draw the steel region of the Fe–Fe₃C phase diagram and make neat sketches of microstructures expected for four compositions between 0.0 and 1.2% C.

- 7.12** From the Fe–Fe₃C phase diagram, for a 0.2% C steel, name the phases and their fractions at equilibrium at the following temperatures:

- (i) just above 1493°C ,
- (ii) just below 1493°C ,
- (iii) just above 725°C , and
- (iv) just below 725°C .

Answer: (i) $f_{\delta} = 0.75$, $f_{\text{L}} = 0.25$ (ii) $f_{\gamma} = 0.94$, $f_{\text{L}} = 0.06$
 (iii) $f_{\alpha} = 0.77$, $f_{\gamma} = 0.23$ (iv) $f_{\alpha} = 0.97$, $f_{\text{Fe}_3\text{C}} = 0.03$.

- 7.13** What is the fraction of proeutectoid cementite in (i) 1.4% C, (ii) 1.0% C, and (iii) 0.7% C steels?

Answer: (i) 0.10, (ii) 0.034, and (iii) 0.

- 7.14** In cooling from the γ range, both proeutectoid ferrite and proeutectoid cementite form at the austenite grain boundaries. If a structure at room temperature contains 5% of grain boundary phase and 95% of pearlite, what are the possible values for the carbon content? Suggest a method to determine the right composition.

Answer: 1.09% and 0.76%.

- 7.15** In the binary phase diagram (Fig. 7.16), mark the various phase fields. Give all the invariant reactions that occur, stating their names.

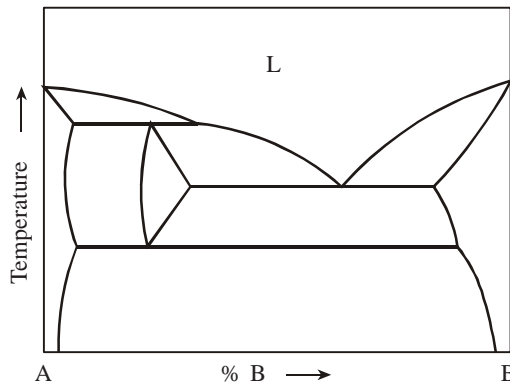


Fig. 7.16 Binary phase diagram of A and B.

7.16 If an impurity raises the melting point of a pure component, how will the zone refining process be affected?

7.17 Derive the degrees of freedom for a system, which has equal number of components and phases.

Answer: 2.

7.18 One solid phase on heating through an invariant temperature becomes two solid phases. Name the invariant reaction. Sketch the phase boundaries near the invariant line.

7.19 In the Pb–Sn system, calculate the alloy composition at which the fraction of total α is $2\frac{1}{2}$ times the fraction of the β phase at 182°C .

Answer: 40.6% Sn.

7.20 For soldering with Pb–Sn alloys, at least 85% of the eutectic mixture is preferred in the microstructure. Determine the composition limits of tin that will satisfy this condition.

Answer: 55.4–67.25%.

7.21 Calculate the fraction of proeutectoid ferrite, eutectoid ferrite and total ferrite in a 0.2% C steel.

Answer: 0.75, 0.22 and 0.97.

7.22 The potassium–sodium binary phase diagram has the following invariant reactions:

	cooling \rightarrow
at 6.9°C ,	$\beta + \text{L} \leftrightarrow \text{Na}_2\text{K}$
wt.% Na	99 47 54

	cooling \rightarrow
at -12.6°C ,	$\text{L} \leftrightarrow \alpha + \text{Na}_2\text{K}$
wt.% Na	23 3 54

(i) Give the name of each of the above reactions.

(ii) Make an approximate sketch of the phase diagram.

(iii) Find the fraction of Na_2K in a 33 wt.% Na alloy at -12.5°C .

Answer: (iii) 0.32.

7.23 In the Ge–P system, a compound GeP exists. At 725°C , Ge, GeP and liquid of 62 at.% P are in three-phase equilibrium. At 577°C , GeP, P and liquid of 94.5 at.% P are in equilibrium. The melting points of Ge and P are 937 and 593°C respectively. Draw an approximate phase diagram of this system. Write down the invariant reactions, giving their names.

7.24 In the Ti–Ti₂Co system, liquid of 27 wt.% Co, β of 17 wt.% Co and Ti₂Co of 38 wt.% Co are in equilibrium at 1020°C . At 685°C , β of 8 wt.% Co, α of 1 wt.% Co and Ti₂Co are in equilibrium. Pure Ti melts at 1670°C and also undergoes a crystal structure change from β (BCC) to α (HCP) on cooling through 882°C . Ti₂Co is stable up to 1058°C .

- (i) Make an approximate sketch of the phase diagram showing temperatures, compositions and phases.
- (ii) Write down the invariant reactions, giving their names.
- (iii) How much α and Ti_2Co are present in a 6 wt.% Co alloy at room temperature?

Answer: (iii) 0.86, 0.14.

7.25 The data on the gold–lead phase diagram are given below. Melting point of Au = 1063°C , Melting point of lead = 327°C . At 434°C , Au, liquid of 43 at.% Pb and Au_2Pb are in equilibrium. At 253°C , Au_2Pb , liquid of 74 at.% Pb and AuPb_2 are in equilibrium. At 222°C , AuPb_2 , AuPb_3 and liquid of 82 at.% Pb are in equilibrium. At 212°C , AuPb_3 , Pb and liquid of 84 at.% Pb are in equilibrium. Draw the phase diagram on a graph paper using a suitable scale. Write down the invariant reactions, giving the name of each.

7.26 In the Fe– Fe_2Nb system, at 1373°C , δ of 3% Nb, liquid of 12% Nb and ε of 27% Nb are in equilibrium. At 1210°C , γ of 1% Nb, δ of 1.5% Nb and ε of 27% Nb are in equilibrium. At 961°C , γ of 0.4% Nb, α of 0.7% Nb and ε of 27% Nb are in equilibrium. All compositions are in atomic%. The melting point of Fe is 1535°C and of Fe_2Nb (33.3 at.% Nb) is 1627°C . The ε phase is stable over a composition range of 27 to 33.3 at.% Nb,

- (i) Draw an approximate phase diagram for this system.
- (ii) Write down the invariant reactions in the system, giving their names.
- (iii) What is the fraction of α in an alloy with 10% Nb at 960°C ?

Answer: (iii) 0.65.

7.27 The Al–Si phase diagram is of the simple eutectic type. At 577°C , solid Al with 1.5% Si dissolved in it, solid Si and liquid of 12.5% Si are in equilibrium. An aluminium wire is welded to a silicon substrate. Just below the weld inside the substrate, the microstructure shows 2% of a eutectic-like mixture. What is the composition at this location?

Answer: 98.28% Si.

MULTIPLE CHOICE QUESTIONS

1. The maximum number of co-existing in a C -component system is
 A. $C - P + 2$ B. $P(C - 1)$ C. $F - C + 2$ D. $C + 2$
2. In a single-component system, the maximum number of phases that can coexist in equilibrium is
 A. 2 B. 3 C. 4 D. 5

3. The degree of freedom when ice, water and water vapour co-exist in equilibrium is
A. 1 B. triple pt C. 0 D. -1
4. The degrees of freedom, when FCC iron and BCC iron co-exist in equilibrium, are
A. 2 B. 1 C. 0 D. -1
5. The phase boundary between alpha and (alpha + beta) regions is called
A. liquidus B. solidus C. solvus D. none of these
6. The reaction that yields two solid phases on cooling a single solid phase is called
A. eutectoid B. peritectoid C. eutectic D. congruent
7. If one solid phase splits into two solid phases on heating, the reaction is
A. eutectic B. peritectic C. eutectoid D. peritectoid
8. The reaction that, on heating one solid phase, yields another solid phase plus one liquid phase is called
A. eutectic B. eutectoid C. peritectic D. peritectoid
9. If alpha of 82% B and liquid of 57% B are in equilibrium in an alloy of 73% B, the fraction of liquid is
A. 0.36 B. 0.64 C. 36% B D. 0
10. If the fraction of liquid with 57% B, which is in equilibrium with solid of 82% B, is 0.7, the overall composition is
A. 0.3 B. 74.5% B C. 64.5% B D. 25% B
11. In the eutectic phase diagram of Ag-Cu system, the solubility limit at 500°C of copper is 3% in the Ag-rich phase and of Ag is 2% in the Cu-rich phase. In sterling silver (92.5% Ag -7.5% Cu), the per cent of copper in the Ag-rich phase at 500°C is
A. 95.26 B. 4.74 C. 3 D. 98
12. The eutectic mixture in a Pb-Sn solder alloy should be 90%. At the eutectic temperature, alpha of 19% Sn, liquid of 62% Sn and beta of 97% Sn are in equilibrium. The possible compositions of the solder alloy are
A. 57.7% Sn B. 61% Sn C. 65.5% Sn D. 66.3% Sn
13. The fraction of pearlite in a 0.55% C steel is
A. 0.55 B. 0.31 C. 0.69 D. 0
14. At 30°C, hot chocolate (liquid) with 35% chocolate and 65% vanilla transforms to chocolate ripple (eutectic mixture of vanilla containing 10% chocolate and chocolate containing 5% vanilla). Just below 30°C, the fraction of chocolate ripple in a composition with 45% chocolate is
A. 0.17 B. 0.83 C. 0.41 D. 0.59

15. Zone refining will be more efficient if the ratio of impurity in the solid to that in the liquid is
A. 0.01 B. 0.1 C. 0.4 D. ~1.0

Answers

- | | | | | |
|-------|----------|-------|-------|-------|
| 1. D | 2. B | 3. C | 4. B | 5. C |
| 6. A | 7. D | 8. C | 9. A | 10. C |
| 11. C | 12. A, C | 13. C | 14. B | 15. A |

Sources for Experimental Data

American Ceramic Society, *Phase Equilibria Diagrams (Phase Diagrams for Ceramists)*, Columbus, Ohio, Vol. I (1964) to Vol. XI (1993).

T.B. Massalski (Ed.), *Binary Alloy Phase Diagrams*, Vols. 1–3, ASM International, Materials Park, Ohio (1990).

Suggestions for Further Reading

W.D. Kingery, H.K. Bowen and D.R. Uhlmann, *Introduction to Ceramics*, Wiley, New York (1976), Chap. 7.

A. Prince, *Alloy Phase Equilibria*, Elsevier, Amsterdam (1966).

F.N. Rhines, *Phase Diagrams in Metallurgy*, McGraw-Hill, New York (1956).

CHAPTER

8

Diffusion in Solids

Diffusion refers to the movement of atoms in solids. Structural control in a solid to achieve the optimum properties is dependent on the rate of diffusion. The carburization of a steel or the oxidation of a metal is controlled by the diffusion rate of atoms (or ions) through the surface layer. The introduction of a very small concentration of an impurity in a solid state device requires knowledge of the diffusion phenomenon.

In this chapter, we first consider the macroscopic laws of diffusion, with a discussion of how they can be used in practical applications. The atomistic model of diffusion is considered in the later sections.

Units

Quantity	SI units		Other units
	Unit	Symbol	
Flux J	mole per square metre per second	$\text{mol m}^{-2} \text{s}^{-1}$	No. of atoms/ cm^2/sec
Concentration c	mole per cubic metre	mol m^{-3}	No. of atoms/ cm^3
Concentration gradient dc/dx	mole per cubic metre per metre	mol m^{-4}	No. of atoms/ cm^4
Diffusion coefficient D	metre squared per second	$\text{m}^2 \text{s}^{-1}$	—
Diffusion constant D_0			
Activation energy for diffusion, Q	kilojoule per mole	kJ mol^{-1}	kcal per mole, eV/atom
Enthalpy of motion of vacancy, ΔH_m			
Enthalpy of formation of vacancy, ΔH_f			
Vibration frequency ν	per second	s^{-1}	—

8.1 Fick's Laws of Diffusion

Diffusion can be defined as the *mass flow process* by which atoms (or molecules) change their positions relative to their neighbours in a given phase under the influence of thermal energy and a gradient. The gradient can be a concentration gradient, an electric or magnetic field gradient or a stress gradient. We shall consider mass flow under concentration gradients only. Thermal energy is necessary for mass flow, as the atoms have to jump from site to site during diffusion. The thermal energy is in the form of the vibrations of atoms about their mean positions in the solid.

We first describe Fick's macroscopic laws of diffusion, before going into the atomistic models. Consider the unidirectional flow of matter in a binary system of A and B atoms. The two types of atoms will move in opposite directions under the influence of a concentration gradient. Let us assume that B is the only moving species. *Fick's First Law* states:

$$\frac{dn}{dt} = -DA \frac{dc}{dx} \quad (8.1)$$

where dn/dt is the number of moles of B atoms crossing per unit time a cross-sectional plane of area A perpendicular to the diffusion direction x and dc/dx is the concentration gradient in the x -direction. D is called the *diffusion coefficient* (or diffusivity) and is a constant characteristic of the system. The diffusion coefficient depends on the nature of the diffusing species, the matrix in which it is diffusing, and the temperature at which diffusion occurs. The negative sign indicates that the flow of matter occurs *down* the concentration gradient. By definition, flux J is flow per unit cross-sectional area per unit time so that Fick's First Law can also be written as

$$J = \frac{1}{A} \frac{dn}{dt} = -D \frac{dc}{dx} \quad (8.2)$$

Fick's First Law can be used to describe flow under *steady state conditions*. It is identical in form to Fourier's law for heat flow under a constant temperature gradient and Ohm's law for current flow under a constant electric field gradient. Under steady state flow, the flux is independent of time and remains the same at any cross-sectional plane along the diffusion direction:

$$J \neq f(x, t) \quad (8.3)$$

Figure 8.1 shows the concentration-distance profile under steady state flow. The profile is a straight line, when D is independent of concentration. When $D = f(c)$, the profile will be such that the product $D(dc/dx)$ is a constant. In neither case, the profile changes with time, under conditions of steady state flow.

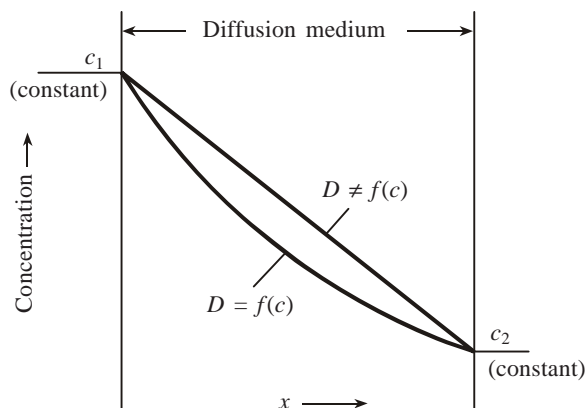


Fig. 8.1 Concentration-distance profile for steady-state flow.

Example 8.1 A steel tank contains hydrogen at a constant pressure of 10 atm, with a vacuum outside. The hydrogen concentration at the inner surface of the tank is equal to 10 kg m^{-3} . The diffusion coefficient of hydrogen in steel at room temperature is $10^{-9} \text{ m}^2 \text{ s}^{-1}$. Calculate the rate at which hydrogen escapes through the wall of the steel tank, which has a thickness of 5 mm.

Solution As soon as steady state flow is established, a constant flux of hydrogen will escape through the tank wall, as the pressure drop inside the tank is negligible. There is vacuum on the outside. So, the concentration of hydrogen on the outer surface of the wall is zero. From Fick's First Law,

$$\text{Flux outward } J = 10^{-9} \times 10 / (5 \times 10^{-3}) = 2 \times 10^{-6} \text{ kg m}^{-2} \text{ s}^{-1}.$$

Fick's Second Law is an extension of the first law to *nonsteady state flow*. Here, at any given instant, the flux is not the same at different cross-sectional planes along the diffusion direction x . Also, at the same cross-section, the flux is not the same at different times. Consequently, the concentration-distance profile changes with time. Nonsteady state flow is frequently met with in practical applications of the diffusion laws.

Consider an elemental slab of thickness Δx along the diffusion distance x . Let the slab cross-section be perpendicular to x and its area be unity. The volume of the slab is then Δx . Under nonsteady state conditions, the flux into the slab J_x is not equal to the flux out of the slab, $J_{x+\Delta x}$. The rate of accumulation (or depletion) of the diffusing atoms within this elemental volume is $(\partial c / \partial t) \Delta x$. It can be expressed as the difference of fluxes in and out of the slab:

$$\left(\frac{\partial c}{\partial t} \right) \Delta x = J_x - J_{x+\Delta x} = J_x - \left\{ J_x + \left(\frac{\partial J}{\partial x} \right) \Delta x \right\} \quad (8.4)$$

or

$$\frac{\partial c}{\partial t} = - \frac{\partial J}{\partial x} \quad (8.5)$$

Equation (8.2) can be rewritten in partial form and substituted in Eq. (8.5):

$$\frac{\partial c}{\partial t} = - \frac{\partial}{\partial x} \left(-D \frac{\partial c}{\partial x} \right) = \partial x \left(D \frac{\partial c}{\partial x} \right) \quad (8.6)$$

Equation (8.6) is Fick's Second Law for unidirectional flow under nonsteady state conditions. If D is independent of concentration, Eq. (8.6) simplifies to

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \quad (8.7)$$

Even though D may vary with concentration, solutions to the differential equation (8.7) are quite commonly used for practical problems, because of their relative simplicity.

8.2 Solution to Fick's Second Law

The solution to Eq. (8.7) for unidirectional diffusion from one medium to another across a common interface is of the general form

$$c(x, t) = A - B \operatorname{erf} \frac{x}{2\sqrt{Dt}} \quad (8.8)$$

where A and B are constants to be determined from the initial and boundary conditions of a particular problem. Here, the diffusion direction x is perpendicular to the common interface. The origin for x is at the interface. The two media are taken to be semi-infinite, that is, only one end of each of them, which is at the interface, is defined. The other end of each is at an infinite distance. The initial uniform concentrations of the diffusing species in the two media are different, with an abrupt change in concentration at the interface.

In Eq. (8.8), 'erf' stands for error function, which is a mathematical function defined as follows:

$$\operatorname{erf} \frac{x}{2\sqrt{Dt}} = \frac{2}{\sqrt{\pi}} \int_0^{x/2\sqrt{Dt}} \exp(-\eta^2) d\eta \quad (8.9)$$

η is an integration variable that gets deleted as the limits of the integral are substituted. The lower limit of the integral is always zero, while the upper limit of the integral is the quantity whose error function is required. The factor $2/\sqrt{\pi}$ is a normalization factor. In Fig. 8.2, $\exp(-\eta^2)$ is plotted against η . The area under the curve from $\eta = 0$ to $x/2\sqrt{Dt}$ is the value of the integral in Eq. (8.9). The area integrated from $\eta = 0$ to $\eta = +\infty$ comes out to be $\sqrt{\pi}/2$ and the area from $\eta = 0$ to $\eta = -\infty$ is $-\sqrt{\pi}/2$.

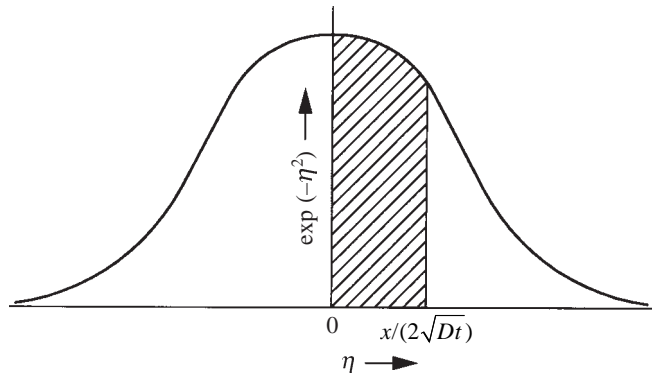


Fig. 8.2 Illustration of the error function. The hatched area is equal to the value of the integral in Eq. (8.9).

Therefore,

$$\text{erf}(\infty) = 2/\sqrt{\pi} \times \sqrt{\pi}/2 = 1$$

$$\text{erf}(-\infty) = 2/\sqrt{\pi} \times -\sqrt{\pi}/2 = -1$$

$$\text{erf}(0) = 0$$

and

$$\text{erf} - \frac{x}{2\sqrt{Dt}} = -\text{erf} \frac{x}{2\sqrt{Dt}}$$

A number of applications of the solution (8.8) are discussed in Sec. 8.3.

8.3 Applications Based on the Second Law Solution

8.3.1 Experimental Determination of D

The diffusion coefficient D can be determined experimentally using a diffusion couple. A diffusion couple, shown in Fig. 8.3, consists of two long bars welded

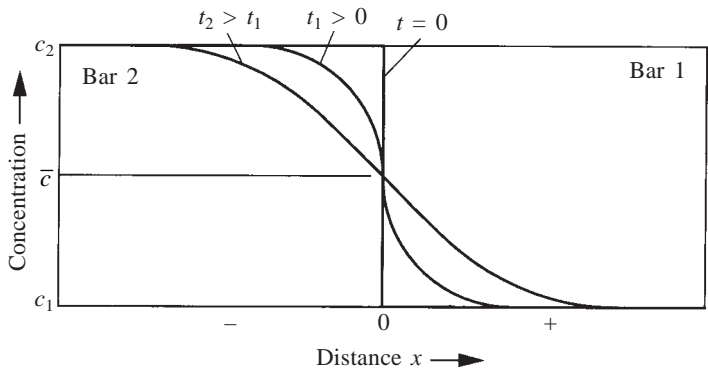


Fig. 8.3 The diffusion couple set-up for nonsteady-state flow.

face to face, the concentration of the diffusing species is one, c_2 , being higher than that in the other, c_1 . Diffusion takes place across the common face. In Fig. 8.3, the concentration is shown along the y -axis; the origin of the diffusion direction x is at the common interface.

The initial conditions for the diffusion couple set-up are:

$$c(x, 0) = \begin{cases} c_1 & x > 0 \\ c_2 & x < 0 \end{cases} \quad (8.10)$$

So, the constants in Eq. (8.8) are given by

$$A = (c_2 + c_1)/2 \quad (8.11)$$

$$B = (c_2 - c_1)/2 \quad (8.12)$$

The thermal energy at room temperature is usually insufficient to cause an appreciable rate of diffusion. If the diffusion couple is heated to a high temperature, say, near the melting point, sufficient thermal energy becomes available and the diffusing species start to move from bar 2 to bar 1. The effect of this movement will be first felt near the common face, and after time t_1 , the concentration distance profile will change near the junction as shown. The concentration has decreased from the initial value of c_2 on the left side of the origin and has increased from the initial value of c_1 on the right side of the origin. After a longer time t_2 at the elevated temperature, the concentration changes penetrate deeper on both sides of the origin, as indicated by the changing concentration-distance profile. It is clear that the flux is changing as a function of x at a given t , as well as with time t at constant x , under the nonsteady state conditions prevailing. When D is independent of concentration, the concentration-distance profiles are symmetrical about the junction as shown in Fig. 8.3. The concentration at the junction is independent of time and is $\bar{c} = (c_1 + c_2)/2$, where \bar{c} is the average concentration.

The diffusion anneal is done at a constant temperature (to keep D constant) for a known length of time t . The temperature chosen is such that there is a measurable amount of diffusion, after a reasonable length of time. After the anneal, thin slices of the couple perpendicular to the diffusion direction x are machined out in a precision lathe and chemically analyzed to give the value of c as a function of x . Since the annealing time t is known, from the initial compositions c_2 and c_1 , A , B and $\text{erf}(x/2\sqrt{Dt})$ can be determined using Eq. (8.8). $(x/2\sqrt{Dt})$ is then obtained from a standard mathematical table of error functions, as shown in Table 8.1. With x and t known, D can be calculated.

TABLE 8.1
The Error Function

z	$\text{erf } (z)$	z	$\text{erf } (z)$
0.000	0.0000	0.85	0.7707
0.025	0.0282	0.90	0.7970
0.05	0.0564	0.95	0.8209
0.10	0.1125	1.0	0.8427
0.15	0.1680	1.1	0.8802
0.20	0.2227	1.2	0.9103
0.25	0.2763	1.3	0.9340
0.30	0.3268	1.4	0.9523
0.35	0.3794	1.5	0.9661
0.40	0.4284	1.6	0.9763
0.45	0.4755	1.7	0.9838
0.50	0.5205	1.8	0.9891
0.55	0.5633	1.9	0.9928
0.60	0.6039	2.0	0.9953
0.65	0.6420	2.2	0.9981
0.70	0.6778	2.4	0.9993
0.75	0.7112	2.6	0.9998
0.80	0.7421	2.8	0.9999

If D is a function of concentration, the above procedure can still be used, by choosing c_2 and c_1 within a narrow composition range for anyone diffusion couple. The experiment can then be repeated with additional diffusion couples to cover the entire composition range.

The diffusion couple experiment can be done at different temperatures of anneal so that D can be determined as a function of temperature. The experimental results indicate an Arrhenius type of relationship:

$$D = D_0 \exp \left(-\frac{Q}{RT} \right) \tag{8.13}$$

where D_0 is a pre-exponential constant and Q is the activation energy for diffusion. The values of D_0 and Q are determined from a plot of $\ln D$ along the y-axis and $1/T$ along the x-axis. This plot yields a straight line, the slope of which is $-Q/R$. The intercept on the y-axis is $\ln D_0$. Table 8.2 lists the experimentally determined values of D_0 and Q for a number of diffusion processes. When a species diffuses in its own matrix, the process is called *self-diffusion*, e.g., copper diffusion in a copper matrix. Some self-diffusion data are also given in the table. These are determined by studying the diffusion of radioactive isotopes of the element in a nonradioactive matrix of the same element.

TABLE 8.2
 D_0 and Q Values for Selected Diffusion Processes*

Diffusion process	$D_0, 10^{-4} \text{ m}^2 \text{ s}^{-1}$	$Q, \text{ kJ mol}^{-1}$
Cu in Cu	0.20	196
Zn in Zn	0.15	94
Al in Al	1.98	143
Fe in Fe (α)	118	281
Ge in Ge	9.3	288
Si in Si	5400	477
W in W	43	640
C in graphite	7	681
H in Fe (α)	0.001	13
N in Fe (α)	0.005	76
C in Fe (α)	0.008	83
C in Fe (γ)	0.7	157
V in Fe	3.9	244
Mn in Fe	4.0	305
Ni in Fe	2.6	295
Zn in Cu	0.73	170
Ni in Cu	2.0	230
Cu in Al	0.25	121

* In many cases, the diffusion coefficient varies significantly with composition of the matrix. Therefore, these values should be considered as approximate. More accurate values can be obtained for specific matrix compositions in the references listed at the end of this chapter.

8.3.2 Corrosion Resistance of Duralumin

Duralumin is primarily an alloy of aluminium with 4% copper. In the properly heat-treated condition, this alloy has a strength which is several times more than that of aluminium. It is, therefore, used widely in the aircraft industry. The corrosion resistance of duralumin is poor compared to that of aluminium. Hence, sheets of duralumin are covered on both sides with thin pure aluminium sheets and passed through rolls to produce a sandwich like material called *Alclad*. When this material is heated to 550°C for giving the appropriate heat treatment to increase the strength, copper diffuses out of duralumin into the outer layers and impairs the corrosion resistance of pure aluminium. To minimize the damage due to diffusion, the thickness of the aluminium sheets used for protection and the time of heating at the elevated temperature should be controlled.

Example 8.2 In Alclad, 20 mm thick duralumin sheets are covered on either side with 0.2 mm thick pure aluminium sheets. For retaining the corrosion resistance, the copper concentration at a depth of 0.1 mm from the outer surface should not exceed 0.4%. How long can the material be kept at 550°C, without damaging the corrosion resistance?

Solution From Table 8.1, the diffusion coefficient of copper in aluminium at 550°C is obtained:

$$\begin{aligned} D_{\text{Cu in Al}} &= D_0 \exp(-Q/RT) \\ &= 0.25 \times 10^{-4} \exp\{-(121.0 \times 10^3)/(8.314 \times 823)\} \\ &= 5.25 \times 10^{-13} \text{ m}^2 \text{ s}^{-1}. \end{aligned}$$

Taking the duralumin-aluminium interface as the diffusion couple interface, we can write

$$c(x, 0) = \begin{cases} 4\% & x > 0 \\ 0\% & x < 0 \end{cases}$$

$$x = 0.2 - 0.1 = 0.1 \text{ mm}$$

$$c(x) = 0.4\%$$

$$D = 5.25 \times 10^{-13} \text{ m}^2 \text{ s}^{-1}$$

$$t = ?$$

From Eq. (8.8),

$$A = 2\%, \quad B = 2\%$$

$$\operatorname{erf}(x/2\sqrt{Dt}) = 0.8$$

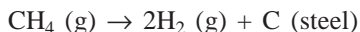
From Table 8.1,

$$x/2\sqrt{Dt} = 0.90$$

$$\begin{aligned} t &= (0.1^2 \times 10^{-6}) / (0.9^2 \times 4 \times 5.25 \times 10^{-13}) \\ &= 5879 \text{ s} \approx 100 \text{ min}. \end{aligned}$$

8.3.3 Carburization of Steel

Surface hardening of steel objects such as gears is frequently done by carburizing or nitriding. The process consists of diffusing carbon (or nitrogen) into the surface layers of the steel object. In pack carburizing, the object is packed in solid carbon powder. In gas carburizing, an atmosphere of methane gas that is rich in carbon surrounds the object to be carburized. Here, the following reaction takes place at the steel surface:



When the steel object is annealed at an elevated temperature in the carburizing medium, carbon diffuses into the steel from the surface under a concentration

gradient. Surface hardening improves the wear resistance of components such as gears, without impairing the bulk mechanical properties such as toughness.

If the carbon content of the carburizing atmosphere remains constant, it would give rise to a constant carbon concentration c_s at the surface of the steel, Fig. 8.4a.

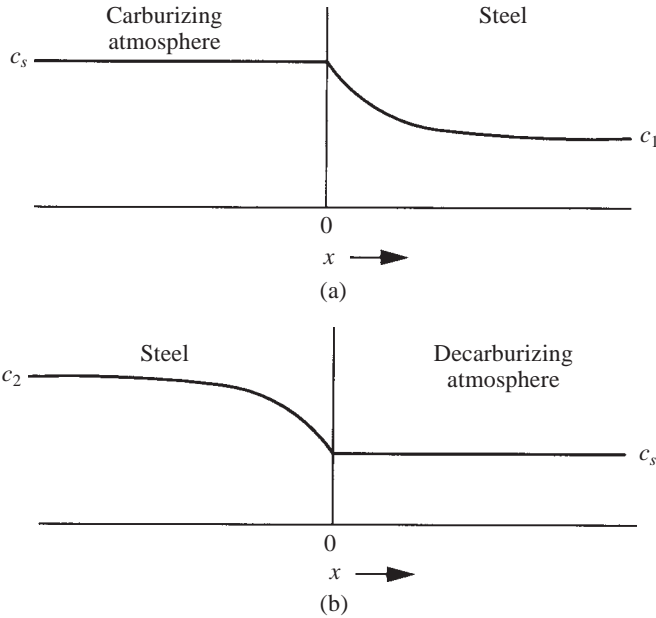


Fig. 8.4 Concentration-distance profiles for (a) carburization, and (b) decarburization.

The initial carbon content of the steel is c_1 . We can then write

$$c(x, 0) = c_1, \quad x > 0$$

$$c(0, t) = c_s$$

So, from Eq. (8.8),

$$A = c_s, \quad B = c_s - c_1$$

With D , c_s and c_1 known, the amount and depth of carbon penetration as a function of time can be computed.

8.3.4 Decarburization of Steel

The opposite of carburization is decarburization. Here, the carbon is lost from the surface layers of the steel, due to an oxidizing atmosphere that reacts with carbon to produce CO or CO₂. The fatigue resistance of steels is lowered due to decarburization and, therefore, it should be avoided by using a protective atmosphere during the heat treatment of the steel. When the heat treatment is carried out in a non-protective atmosphere (e.g., air), the extent of decarburization

can be estimated from the diffusion equation and post-machining operations can be undertaken to remove the decarburized layer.

Referring to Fig. 8.4b, the decarburizing atmosphere is equivalent to a lower carbon content at the steel surface, as compared to the initial carbon concentration (c_2) of the steel. Here

$$c(x, 0) = c_2, \quad x < 0$$

$$c(0, t) = c_s$$

So, from Eq. (8.8),

$$A = c_s, \quad B = c_2 - c_s$$

With D , c_s and c_2 known, the depth and the degree of decarburization can be determined at a function of the heat treating time t .

Example 8.3 At 950°C, a 0.8% carbon steel is getting decarburized for a duration of 4 hr in an atmosphere equivalent to 0% carbon at the surface of the steel. Determine the minimum depth up to which post-machining is to be done, if the carbon content at the surface after-machining should not be below 0.6%.

Solution From Table 8.2,

$$D_{\text{C in Fe}(\gamma)} = 0.7 \times 10^{-4} \exp \{(-157 \times 1000)/(8.314 \times 1223)\}$$

$$= 1.38 \times 10^{-11} \text{ m}^2 \text{ s}^{-1}$$

$$c_2 = 0.8\%$$

$$c_s = 0\%$$

$$c(x) = 0.6\%$$

$$t = 4 \times 3600 = 14\,400 \text{ s}$$

$$D = 1.38 \times 10^{-11} \text{ m}^2 \text{ s}^{-1}$$

$$x = ?$$

Substituting in Eq. (8.8), we get

$$\operatorname{erf} \frac{x}{2\sqrt{Dt}} = -0.75$$

From Table 8.1,

$$\frac{x}{2\sqrt{Dt}} = -0.81$$

$$x = -7.22 \times 10^{-4} \text{ m}$$

The depth up to which machining is required is $\sim 0.75 \text{ mm}$.

8.3.5 Doping of Semiconductors

Semiconductor devices are doped with small controlled quantities of impurities for obtaining the desired electrical characteristics. The dopant atoms may be diffused into the pure semiconductor crystal from a gaseous atmosphere. The depth of penetration and the amount of dopant in the crystal can be estimated following the same procedure as given for the carburization problem.

Example 8.4 To produce a *p*-type semiconductor, the third column element boron is doped in pure silicon. The doping is done through a B_2O_3 vapour phase of partial pressure equal to 1.5 N m^{-2} . This atmosphere is equivalent to a surface concentration of 3×10^{26} boron atoms per m^3 . Calculate the time required to get a boron content of 10^{23} atoms per m^3 at a depth of $2 \text{ }\mu\text{m}$. The doping temperature is 1100°C and $D_{B \text{ in Si}}$ at this temperature is $4 \times 10^{-17} \text{ m}^2 \text{ s}^{-1}$.

Solution

$$c_1 = 0$$

$$c_s = 3 \times 10^{26} \text{ atoms}$$

$$c(x) = 10^{23} \text{ atoms}$$

$$x = 2 \times 10^{-6} \text{ m}$$

$$D = 4 \times 10^{-17} \text{ m}^2 \text{ s}^{-1}$$

$$t = ?$$

Substituting in Eq. (8.8), we get

$$\text{erf} (x/2\sqrt{Dt}) = 0.99967$$

From Table 8.1,

$$(x/2\sqrt{Dt}) = 2.55$$

$$\begin{aligned} t &= (2^2 \times 10^{-12}) / (4 \times 2.55^2 \times 4 \times 10^{-17}) \\ &= 3845 \text{ s} \end{aligned}$$

When accurate calculations are not needed, the depth of penetration of the diffusing species can be quickly estimated, using the following approximation:

$$x = \sqrt{Dt} \quad (8.14)$$

If c_1 is negligible compared to c_2 , for the diffusion couple set-up, this approximation yields the concentration at x as $c(x) = 0.24 c_2$.

8.4 The Kirkendall Effect

In a binary solution of A and B, the rates at which A and B diffuse are not necessarily equal. Usually, the lower melting component diffuses much faster than the other. This leads to certain interesting effects as first observed by Kirkendall.

Inert markers (thin rods of a high melting point substance which is insoluble in the diffusion matrix) are placed at the weld joint of the couple, prior to the diffusion anneal. These markers are found to shift during the anneal in the same direction as the slower moving species. The extent of this shift is found to be proportional to the square root of the diffusion time. This kind of movement indicates that the net mass flow due to the difference in diffusivities is being compensated by a bulk flow of matter in the opposite direction within the diffusion zone. That is, lattice planes are created on one side of the diffusion zone, while they are destroyed on the other side, and the resulting bulk flow carries the markers along. Notice that the bulk flow occurs relative to the ends of the diffusion couple. It is quite a different phenomenon from the diffusion process itself. In many cases, porosity is observed on the lower-melting component side, indicating that the bulk flow does not fully compensate for the difference in diffusivities of the two species.

The following analogy of gaseous interdiffusion aids in the understanding of the Kirkendall effect. Let hydrogen and argon at the same pressure be kept in two chambers interconnected through a tube. A frictionless piston in the tube separates the gases. On opening an orifice in the piston, the gases interdiffuse. The lighter hydrogen will diffuse faster, resulting in a pressure difference that will tend to shift the piston in the same direction as the slower diffusing argon is moving.

8.5 The Atomic Model of Diffusion

Diffusion occurs as a result of repeated jumps of atoms from their sites to other neighbouring sites. Even when atoms jump randomly, a net mass flow can occur down a concentration gradient, when a large number of such jumps take place. The unit step in the diffusion process is a single jump by the diffusing species.

In *interstitial diffusion*, solute atoms which are small enough to occupy interstitial sites, diffuse by jumping from one interstitial site to another, as illustrated in Fig. 8.5a. In *vacancy diffusion*, atoms diffuse by interchanging

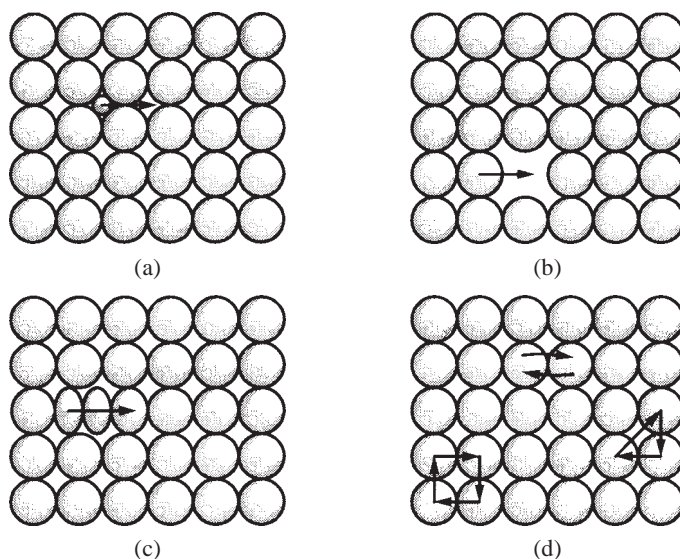


Fig. 8.5 Mechanisms of diffusion.

positions with neighbouring vacant sites, Fig. 8.5b. In interstitialcy mechanism, the configuration where two atoms share a common atomic site moves through the crystal, Fig. 8.5c. In the direct interchange mechanism, Fig. 8.5d, severe local distortion is involved during motion. In the ring mechanism, three or four atoms in the form of a ring move simultaneously round the ring, thereby interchanging their positions, Fig. 8.5d. We will consider in some detail only the interstitial and the vacancy diffusion.

The *unit step in interstitial diffusion* is the jump of the diffusing atom from one interstitial site to a neighbouring site. If the interstitial solution is dilute, the probability that the neighbouring site will be vacant is high. Consider the diffusion of interstitial carbon in FCC iron. In a 1% carbon steel, the atomic fraction of carbon is about 0.05. The probability that an interstitial site will be vacant and be available for a diffusing atom to jump into it is then $1 - 0.05 = 0.95$, which can be approximated to unity.

The potential energy of the crystal as a function of the position of the diffusing interstitial is schematically shown in Fig. 8.6. The energy is a

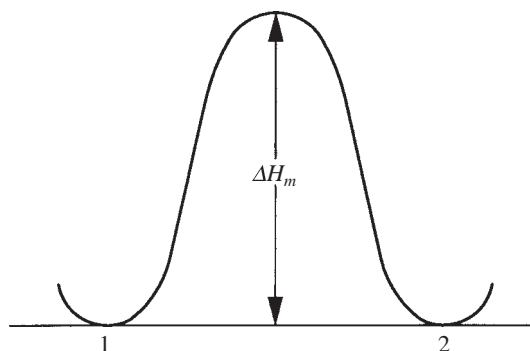


Fig. 8.6 Variation in potential energy along the path of the diffusing atom.

minimum when the interstitial atom is at site 1 or at site 2. The energy increases along the path from site 1 to 2, reaching a maximum at the midpoint of the path. The interstitial atom has to push its way through the parent atoms, with its outer electron cloud having a maximum overlap with the parent atoms at the midpoint. The potential energy barrier ΔH_m is called *the enthalpy of motion*. It is clear that the unit step involves a momentary increase in the enthalpy of the crystal and that thermal energy is needed to overcome the barrier. At absolute zero, the probability of a jump occurring is zero. As the temperature increases, the vibrational amplitude of atoms also increases. There is a statistical distribution of vibrational amplitudes and the probability that an atom will have sufficient vibrational energy to overcome the energy barrier is $\exp(-\Delta H_m/RT)$ from the Maxwell-Boltzmann statistics. If ν is the frequency of the vibrations and ν' is the number of successful jumps per unit time, we have

$$\nu' = \nu \exp(-\Delta H_m/RT) \quad (8.15)$$

v' is directly related to the diffusion coefficient D . A simplified view of this relationship is as follows: For an individual jump, the composition varies from unity to zero over the jump distance δ which is about one atomic diameter, see Fig. 8.7.

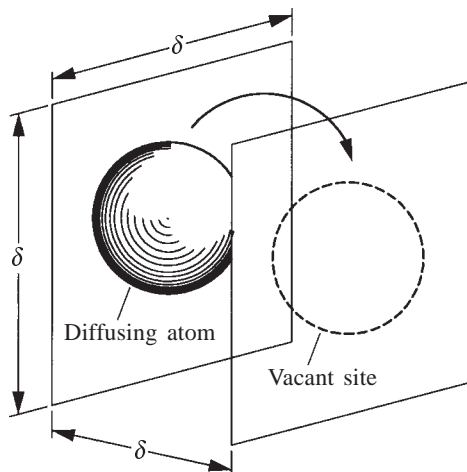


Fig. 8.7 The unit jump.

This composition change occurs in a volume of δ^3 . We can then approximate the concentration gradient dc/dx to $(-1/\delta^3)/\delta = -1/\delta^4$, where concentration is expressed as number of atoms per unit volume. During an individual jump, the flow is across a cross-sectional area δ^2 . So, the flux is v'/δ^2 , in units of number of atoms per unit area per second. The diffusion coefficient D is given by

$$D = \frac{J}{-(dc/dx)} = \frac{v'}{\delta^2} \delta^4 = v' \delta^2$$

$$= v \delta^2 \exp(-\Delta H_m/RT) \quad (8.16)$$

Comparison of Eqs. (8.16) and (8.13) shows that the pre-exponential constant $D_0 = v \delta^2$ and that the activation energy for interstitial diffusion $Q = \Delta H_m$.

Substitutional diffusion generally proceeds by the vacancy mechanism. In the unit step, a substitutional atom jumps into a neighbouring vacant site. Here, we have to take into account the probability that the neighbouring site will be vacant. In Chap. 6, the number of vacancies in thermal equilibrium in a crystal was shown to be $n/N = \exp(-\Delta H_f/RT)$, where ΔH_f is the *enthalpy of formation* of vacancies. The fraction $\exp(-\Delta H_f/RT)$ also gives the probability that a given atomic site will be vacant. In addition, we have to consider as for the interstitial case the probability that the vibrating substitutional atom will have sufficient energy to push its way through to the vacant site. This probability is equal to $\exp(-\Delta H_m/RT)$, where ΔH_m is the *enthalpy of motion* as applicable to the jump

of the substitutional atom. *The product of these two probabilities gives the probability that a vacant site will be available at the same time as a vibration of sufficient amplitude occurs.* The number of successful jumps per unit time is then given by

$$v' = v \exp(-\Delta H_m/RT) \exp(-\Delta H_f/RT) \quad (8.17)$$

As in the interstitial case, the diffusion coefficient is related to v' , yielding

$$D = v\delta^2 \exp\left(-\frac{\Delta H_m + \Delta H_f}{RT}\right) \quad (8.18)$$

The pre-exponential constant of Eq. (8.13) is equated to $v\delta^2$ and the activation energy for diffusion $Q = \Delta H_m + \Delta H_f$. Table 8.3 compares the values of ΔH_f and ΔH_m with the experimental activation energy Q for self-diffusion in silver and gold occurring by the vacancy mechanism.

TABLE 8.3
Calculated and Experimental Activation Energies for Vacancy Diffusion
(kJ mol⁻¹)

Element	ΔH_f	ΔH_m	$\Delta H_m + \Delta H_f$	Q
Silver	97	80	177	174
Gold	95	79	174	184

There is no probability factor involving ΔH_f for interstitial diffusion, as we have already noted. This fact is responsible for interstitial diffusion generally being much faster than substitutional diffusion by the vacancy mechanism. For example, the diffusion coefficient of carbon in FCC iron at 1000°C is $3 \times 10^{-11} \text{ m}^2 \text{ s}^{-1}$, while that of nickel in FCC iron at the same temperature is much less, $2 \times 10^{-16} \text{ m}^2 \text{ s}^{-1}$.

8.6 Other Diffusion Processes

In ionic crystals, *Schottky and Frenkel defects* assist the diffusion process. When Frenkel defects dominate in an ionic crystal, the cation interstitial of the Frenkel defect carries the diffusion flux. When Schottky defects dominate, the cation vacancy carries the diffusion flux. In addition to these defects in thermal equilibrium, ionic crystals may have defects generated by impurities and by deviation from stoichiometry, see Sec. 6.1. For example, a Cd^{2+} cation in a NaCl crystal generates a cation vacancy. Since cation vacancies carry the diffusion flux in NaCl, the presence of cadmium as an impurity increases the diffusivity of NaCl. Even a small concentration such as 0.1% will markedly increase the diffusivity, in relation to that of pure NaCl, as the number of cation vacancies generated by this concentration of the impurity is several orders of magnitude larger than that in thermal equilibrium.

The generation of point defects due to deviation from stoichiometry was discussed in Chap. 6. Referring to Fig. 6.3, excess Zn^{2+} interstitials are present in a non-stoichiometric compound of ZnO . These increase the diffusivity of Zn^{2+} ions markedly. Similarly, the presence of cation vacancies in a nonstoichiometric FeO increases the diffusivity of the cations in FeO .

Measurements of electrical conductivity of ionic crystals yield diffusivity from the *Nernst-Einstein relation*:

$$\frac{\sigma}{D} = \frac{ne^2z}{kT} \quad (8.19)$$

where σ is the electrical conductivity due to the diffusing ions, n is the number of diffusing ions per unit volume, z their valency and e is the electronic charge. This electrical conductivity is called *ionic conductivity*, in contrast to the more familiar form of electrical conduction by the flow of electrons. It is not surprising that this electrical conductivity and the diffusivity are linearly related, as both of them arise from the same atomic process, viz., the diffusion of ionic defects.

Up till now, we have considered diffusion in the crystal with the aid of point imperfections. This can be called lattice diffusion. The diffusion process may also be aided by other imperfections in the crystal, such as line and surface imperfections. Consider the external surface of a crystal. Here, there are no constraints on one side. Hence, the enthalpy of motion for an atom diffusing *along* the surface is considerably less than that for an atom moving *within* the crystal. In a polycrystalline material, the grain boundary regions are not as closely packed as the crystal. ΔH_m for grain boundary diffusion is correspondingly smaller than that for lattice diffusion. The experimentally determined activation energy Q shows the following qualitative variation:

$$Q_{\text{surface}} < Q_{\text{grain boundary}} < Q_{\text{lattice}} \quad (8.20)$$

Likewise, diffusion along dislocation lines known as *pipe diffusion* is faster than lattice diffusion.

A lower activation energy Q does not necessarily mean that diffusion along high diffusivity paths such as grain boundaries and external surfaces will always dominate the lattice diffusion. The cross-sectional area across which mass transport can occur is usually much smaller for special diffusion paths than for lattice diffusion.

Example 8.5 In a cylindrical crystal of radius r ($r = 10$ mm), calculate the ratio of cross-sectional area available for diffusion through the surface layers to the area available for mass transport through the cylinder.

Solution The cross-sectional area for diffusion along the axis of the cylinder is

$$\pi r^2 = 314 \text{ mm}^2$$

Assuming the effective thickness of the surface to be 4 \AA (about two atomic diameters), the cross-sectional area for diffusion along the surface is

$$2\pi r \times 4 \times 10^{-7} = 2.51 \times 10^{-5} \text{ mm}^2$$

The ratio of the two cross-sectional areas is given by

$$2.51 \times 10^{-5} / 314 = 8 \times 10^{-8}$$

The diffusion coefficient for self-diffusion of silver determined for a single crystal and a polycrystalline solid are shown in Fig. 8.8, as a function of $1/T$. At

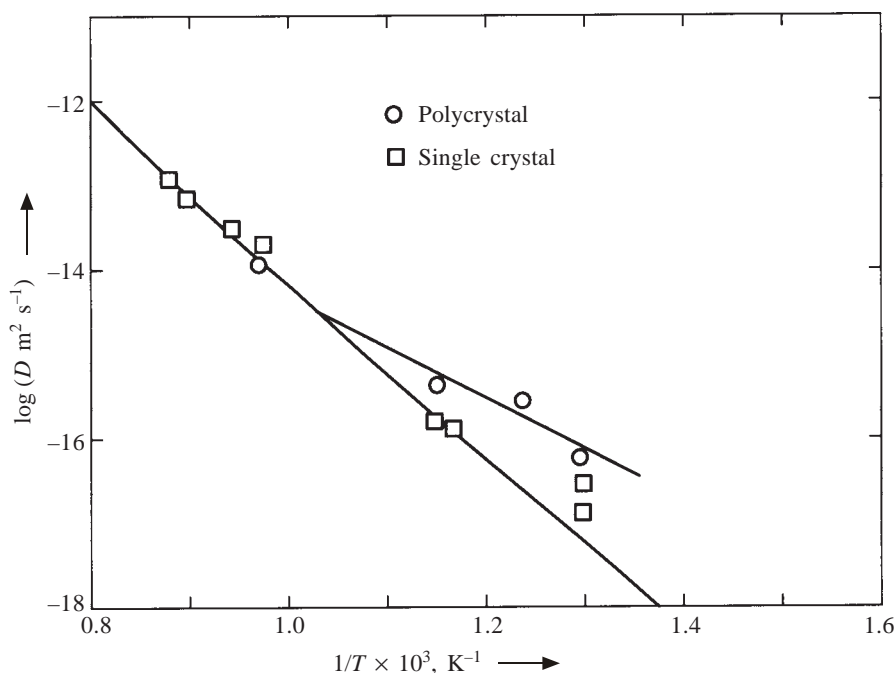


Fig. 8.8 Diffusion data for a single crystal and for polycrystalline silver.

lower temperatures, diffusion through the polycrystalline material is faster, as grain boundary diffusion dominates here. At higher temperatures, lattice diffusion becomes dominant so that the diffusivity is the same for both single crystal and polycrystalline silver. The activation energies calculated from the slopes yield

$$Q_{\text{gb}} = 110 \text{ kJ mol}^{-1}$$

$$Q_{\text{lattice}} = 192 \text{ kJ mol}^{-1}$$

Gaseous molecules diffuse through long chain polymers, without much chemical interaction (bonding) with the chains, using the void space between them. Smaller molecules such as H_2 diffuse more rapidly than larger molecules. The segmental mobility in a polymer assists the diffusion process, as half of a

rotation of a chain segment can be equivalent to the diffusing molecule jumping from one side to the other side of the chain. Cross-linking which restricts chain mobility is found to decrease the diffusion rate. Adding a plasticizer, which allows greater chain mobility, increases the diffusion rate.

Gaseous molecules such as H_2 or He can easily diffuse through the open network structure of silicate glasses. There have been instances, where a liquid helium plant near a high-vacuum equipment built of glass had posed problems in obtaining the desired vacuum. The higher concentration of helium in the surrounding atmosphere finds its way inside the vacuum equipment through the glass!

SUMMARY

1. Fick's First Law is applicable under steady state conditions of mass flow. Here, the flux is proportional to the concentration gradient. The proportionality factor is the diffusion coefficient D .
2. Under nonsteady-state conditions, the flux changes with time and position along the diffusion direction. Fick's Second Law describes the nonsteady-state flow.
3. Solutions to Fick's Second Law are available for a given set of initial and boundary conditions. The solution to the diffusion couple set-up can be used to determine D experimentally. It can be adapted for other practical problems such as doping of a semiconductor crystal and carbon diffusion in steel. Experimental values of D indicate an Arrhenius type of relationship between D and T .
4. Substitutional solutes diffuse by interchanging positions with vacancies. Interstitial solutes simply jump to the next interstitial site, without having to wait for a vacancy. Consequently, interstitial diffusion is much faster than substitutional diffusion.
5. The presence of impurities and deviations from stoichiometry enhance the diffusion rate of ions in ionic crystals by several orders of magnitude.
6. The activation energy for diffusion along surfaces and grain boundaries is lower than that for lattice diffusion. Diffusion along these special paths is effective at low temperatures or for very small particle or grain sizes, which have a high boundary area to volume ratio.

PROBLEMS

- 8.1** The ratio of diffusion rate of silver in silicon at 1350°C to that at 1100°C was found to be 8 in a doping process. Calculate the activation energy Q for silver diffusion in silicon.

Answer: 154 kJ mol^{-1} .

- 8.2** The diffusion rate of A in B was studied at 500°C and 850°C. It was found that, for the same diffusion time, the depths of penetration x_1 and x_2 in the two experiments were in the ratio of 1 : 4. Calculate the activation energy for diffusion of A in B.

Answer: 57 kJ mol⁻¹.

- 8.3** Amorphous selenium used as a semiconductor material exhibits unusual diffusion characteristics. The following is a set of experimental data for self-diffusion in amorphous selenium. Calculate D_0 and Q and comment on your results.

$T, ^\circ\text{C}$	$D, \text{m}^2 \text{s}^{-1}$
35	7.7×10^{-16}
40	2.4×10^{-15}
46	3.2×10^{-14}
56	3.2×10^{-13}

Answer: $D_0 = 10^{27} \text{ m}^2 \text{s}^{-1}$; $Q = 248 \text{ kJ mol}^{-1}$.

- 8.4** At 900°C, what is the time required to carburize a steel with an initial composition of 0.2% carbon to 1% carbon at a depth of 0.2 mm? Assume a constant surface concentration of 1.4% carbon due to the carburizing atmosphere.

Answer: 14 900 s.

- 8.5** A 1.2% carbon steel is getting decarburized in an atmosphere of 0.0% carbon. After some time t , plot (i) $c(x)$ curve near the surface of the steel, and (ii) $J(x)$ curve below the above curve, using the same x -axis.

- 8.6** A diffusion couple of 95% Cu-5% Zn and pure copper is annealed at 900°C for 50 hr. The zinc concentration at a depth of 2 mm inside the copper bar was found to be 0.3% after the anneal. Determine the diffusion coefficient of zinc in copper.

Answer: $5.0 \times 10^{-12} \text{ m}^2 \text{s}^{-1}$.

- 8.7** In a steel, during carburization at 937°C, 0.6% carbon is found at a depth of 0.2 mm after 1 hr. Find the time required to achieve the same concentration at the same depth, if carburization is done at 1047°C.

Answer: 980 s.

- 8.8** Compare the diffusivities of hydrogen, nitrogen, and nickel in iron at 300 K and explain the difference between the three values.

- 8.9** Compute the rate at which a vacancy jumps in copper at 20°C. The activation barrier for the jump is 100 kJ mol⁻¹.

Answer: $1.5 \times 10^{-5} \text{ s}^{-1}$.

- 8.10** An Al-4% Cu alloy is heated to 550°C during heat treatment and quenched to room temperature. Immediately after quench, the diffusion rate of

copper (which proceeds by a vacancy mechanism) was found to be 10^7 times faster than what would be expected from the listed diffusion data. What fraction of vacancies in equilibrium at 550°C is retained at room temperature by the rapid quenching? The enthalpy of motion of vacancy in this alloy is 50 kJ mol^{-1} .

Answer: 0.14.

- 8.11** How will the diffusivity of NaCl change, when it is doped with (i) KCl, and (ii) CaCl_2 ? Explain.

- 8.12** Determine the ratio of cross-sectional area available for diffusion along the surface and through the lattice, when the two diffusion rates are equal at room temperature. Assume D_0 to be the same for both the processes. $Q_{\text{surface}} = 100 \text{ kJ mol}^{-1}$ and $Q_{\text{lattice}} = 150 \text{ kJ mol}^{-1}$.

Answer: 2×10^{-9} .

- 8.13** Find the grain size of a polycrystalline solid for the same amount of material to be transported through (i) the grain and (ii) the grain boundary at 500°C . Assume that the grains are cube shaped and the grain boundaries are 5 \AA thick.

For lattice diffusion: $D_0 = 0.7 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$
 $Q = 188 \text{ kJ mol}^{-1}$

For grain boundary diffusion: $D_0 = 0.09 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$
 $Q = 90 \text{ kJ mol}^{-1}$

Answer: 0.54 mm.

- 8.14** Make a plot of the activation energy Q for diffusion of different species as a function of the melting point of the species. Comment on your result.

- 8.15** From the data in Table 8.2, calculate the diffusion coefficient of carbon in ferrite (α) and austenite (γ) at 900°C . Explain the difference in the values on the basis of the two crystal structures.

Answer: $1.61 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$ in α ; $7.14 \times 10^{-12} \text{ m}^2 \text{ s}^{-1}$ in γ .

- 8.16** In a diffusion anneal, to what level should the initial temperature of 900°C be increased to double the depth of penetration? $D_0 = 0.4 \text{ m}^2 \text{ s}^{-1}$ and $Q = 100 \text{ kJ mol}^{-1}$.

Answer: 1083°C .

- 8.17** An amount Q of a dopant is deposited on the surface of a silicon substrate. During a subsequent anneal without the dopant in the atmosphere, the concentration c of the dopant as a function of depth x and time t is given by

$$c = (Q/\sqrt{\pi DT}) \exp [-x^2/(4Dt)]$$

Show that this is a solution of Fick's Second Law, when D is independent of concentration.

- 8.18** A steel containing 0.002% N is to be nitrided to yield a nitrogen content of 0.12% at depth of 4 mm from the surface. The nitriding atmosphere is equivalent to a surface concentration of 0.35% N. How long is to be the nitriding process? The steel is BCC (α) at the nitriding temperature of 700°C.

Answer: 53 hr

MULTIPLE CHOICE QUESTIONS

- The unit of the diffusion coefficient D is
A. m s^{-2} B. $\text{m}^{-2} \text{s}^{-1}$ C. $\text{m}^2 \text{s}^{-1}$ D. $\text{m}^2 \text{s}$
- The unit of flux J is
A. $\text{atoms m}^{-2} \text{s}^{-1}$ B. $\text{atoms m}^2 \text{s}^{-1}$
C. $\text{moles m}^{-2} \text{s}^{-1}$ D. $\text{moles m}^{-3} \text{s}^{-1}$
- The error function of ∞ is
A. 0 B. -1 C. 1 D. ∞
- The error function of 0 is
A. -1 B. 0 C. 1 D. ∞
- For the same diffusion time, the depth of diffusion penetration at 500 and 850°C is in the ratio of 1 : 6. The activation energy for diffusion is
A. 57 kJ mol^{-1} B. 37 kJ mol^{-1}
C. 114 kJ mol^{-1} D. 74 kJ mol^{-1}
- If $D_0 = 0.4 \times 10^{-4} \text{ m}^2 \text{s}^{-1}$ and $Q = 100 \text{ kJ mol}^{-1}$, to double the depth of penetration, the initial temperature of 900°C should be increased to
A. 910°C B. 923°C C. 986°C D. 1083°C
- In a steel, during carburization at 937°C, 0.6% carbon is found at a depth of 0.2 mm after 1 hr. The time required to get 0.6% C at double this depth at the same temperature is
A. 60 s B. 1.414 hr C. 2 hr D. 4 hr
- Among the following elements, the one with the largest diffusion coefficient in steel at 1000°C is
A. Mn B. W C. Ni D. C
- The fastest diffusing species in Fe is
A. H B. Ni C. C D. W
- If the diffusion jump distance is 1.5 Å, the theoretical value of D_0 in $\text{m}^2 \text{s}^{-1}$ is approximately
A. 1.5×10^3 B. 1.5×10^{-3} C. 2.25×10^{-7} D. 2.25×10^7

11. If the enthalpy of motion of a vacancy is 100 kJ mol^{-1} , the time that a vacancy takes to jump to an adjacent site at 25°C is about
 A. $3 \times 10^{17} \text{ s}$ B. $2 \times 10^{26} \text{ s}$ C. 10^{13} s D. $3 \times 10^4 \text{ s}$
12. The units of the ratio σ/D ($= ne^2z/kT$) are
 A. $\text{A}^2 \text{ m}^{-3} \text{ J}^{-1}$ B. $\text{C}^2 \text{ m}^{-3} \text{ J}^{-1}$ C. $\text{C}^2 \text{ m}^{-3} \text{ J}$ D. $\text{V}^2 \text{ A}^{-2} \text{ m}^{-3} \text{ J}$

Answers

- | | | | | |
|-------|---------|------|------|-------|
| 1. C | 2. A, C | 3. C | 4. B | 5. D |
| 6. D | 7. D | 8. D | 9. A | 10. C |
| 11. D | 12. B | | | |

Source for Experimental Data

E.A. Brandes, *Smithells Metals Reference Book*, Butterworths, London (1983).

Suggestions for Further Reading

- J. Crank and G.S. Park, *Diffusion in Polymers*, Academic Press, New York (1968).
- W.D. Kingery, H.K. Bowen and D.R. Uhlmann, *Introduction to Ceramics*, Wiley, New York (1976), Chap. 6.
- P.G. Shewmon, *Diffusion in Solids*, McGraw-Hill, New York (1963).

The phases present in a material may undergo changes, as a function of temperature or pressure. Such changes are called **phase transformations**. The control of structure and thereby the properties of materials is often achieved by deliberately inducing or suppressing a phase transformation. A knowledge of the kinetics and the mechanism of the transformation is essential for this purpose. In this chapter, we shall discuss the elementary principles that govern phase transformations in materials. In the latter half of the chapter, a number of applications of practical interest based on these principles are discussed.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Free energy change Δg } Enthalpy change Δh }	joule per cubic metre	J m^{-3}	{ erg/cm ³ , cal/cm ³
Energy of nucleation Δf^*	joule per event or kilojoule per mole	J or kJ mol^{-1}	erg, eV
Interfacial energy γ	joule per square metre	J m^{-2}	erg/cm ²
Strain energy ε } Strain energy factor A }	joule per cubic metre	J m^{-3}	erg/cm ³
Contact angle θ	radian	rad	°(degree)
Enthalpy of diffusion ΔH_d	kilojoule per mole	kJ mol^{-1}	kcal/mole
Vibration frequency ν or ν'	per second	s^{-1}	—
Nucleation rate, dN/dt	per cubic metre per second	$\text{m}^{-3} \text{s}^{-1}$	—
Growth rate, U ($= dr/dt$)	metre per second	m s^{-1}	—

(Cont.)

Units (cont.)

Quantity	SI units		Other units
	Unit	Symbol	
Transformation rate, dX/dt	per second	s^{-1}	—
Viscosity η	pascal second	Pa s	poise
Tensile strength	meganewton per square metre	$MN\ m^{-2}$	kgf/mm ² , psi
Hardness R_c	—	—	—
Specific volume	cubic metre per kilogram	$m^3\ kg^{-1}$	ft ³ /lb, cm ³ /gm

Constants

Boltzmann's constant k	$= 1.38 \times 10^{-23}\ J\ K^{-1}$
Gas constant R	$= 8.314\ J\ mol^{-1}\ K^{-1}$

9.1 Time Scale for Phase Changes

The time taken for a transformation to go to completion is important in the control of the structure of a material. On this depends our ability to suppress or induce a phase transformation. The time taken depends on the nature and the mechanism by which a phase transformation is brought about. The free energy change during the transformation also determines its rate. Consider a simple transformation where a liquid L transforms to crystals of β :



The free energies of the two phases are shown in Fig. 9.1. At the equilibrium melting point T_m , the free energies are equal. Above the melting point, the liquid

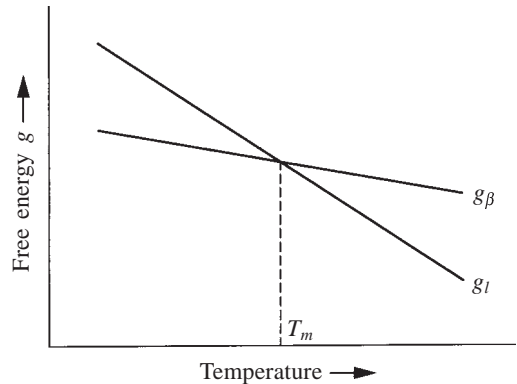


Fig. 9.1 Schematic variation of the free energy g of a material in the liquid and crystalline forms as a function of temperature.

has a lower free energy than β , so that the free energy change for the transformation $L \rightarrow \beta$ is positive. The reaction cannot proceed under such conditions. Below T_m , the free energy of the β -phase is lower and the free energy change Δg for the above reaction is negative:

$$\Delta g = g_\beta - g_L < 0 \quad (9.2)$$

The transformation can now proceed spontaneously. Just below the equilibrium temperature, where the magnitude of the free energy change is very small, the reaction proceeds very slowly. With a large supercooling below T_m , the magnitude of the free energy change is large and the reaction can proceed much faster.

A qualitative time scale for phase transformations is indicated in Table 9.1. The following general discussion is valid for a phase transformation that occurs on cooling a material from an elevated temperature such as the $L \rightarrow \beta$ reaction.

TABLE 9.1
Time Scale for Phase Transformations

Qualitative nature	Time for transformation	Suppressibility
Extremely fast	microseconds	Insuppressible
Very fast	milliseconds to seconds	Suppressed by very fast cooling
Normal	seconds to hours	Suppressible
Slow	hours to days	Suppressed easily
Extremely slow	years	Transformation virtually impossible

Extremely fast transformations, having time scale in microseconds, cannot be normally suppressed. However, in recent years, a very fast cooling technique called *splat cooling* has been developed, which yields cooling rates exceeding a million degrees per second. With such fast cooling rates, even those transformations, which are traditionally considered insuppressible such as the crystallization of a liquid metal, can be suppressed. *Metallic glass* produced by splat cooling has several unusual properties.

Very fast transformations going to completion in a fraction of a second can be suppressed by drastic quenching. Whether drastic quenching is feasible in industrial practice or not, will depend on the size and the thermal conductivity of the material and the effect of thermal stresses arising from the temperature difference between the inside and the outside of the component.

Normal transformations that take place in a matter of seconds or minutes are usually suppressible. There is no problem in suppressing a slow transformation that takes hours to complete. The problem here may turn out to be the other way. If the transformation is desired, we may want to accelerate it so that the industrial production time is cut. Very slow transformations taking years to set in are usually ignored, as the time required for any appreciable amount of transformation can easily exceed the service life of the component in question.

NUCLEATION AND GROWTH

Once the condition for the spontaneous occurrence of a phase transformation has been satisfied, that is, $\Delta g < 0$, the next question is by what mechanism the transformation takes place. Consider the formation of β crystals from a liquid. It is highly improbable that all the atoms in the liquid can coordinate their movements and, at the same instant of time, change over to the crystalline configuration of β . The more likely event is the initial formation of the crystalline configuration among *a few atoms of the liquid*. This arrangement then *extends* into the surrounding liquid, by more and more atoms leaving the liquid and joining the crystalline configuration. The transformation thus proceeds *step by step*, as illustrated in Fig. 9.2, with a number of crystals growing from different points. The growing β crystals eventually impinge on one another and stop growing to yield a polycrystalline solid.

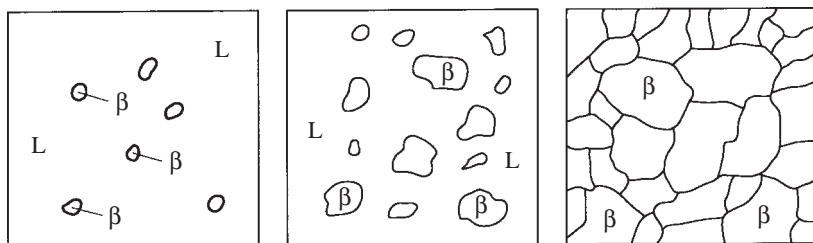


Fig. 9.2 Progressive transformation of the liquid to β crystals by nucleation and growth.

The transformation can be divided into two discrete steps that occur one after the other:

- (i) the formation of tiny stable particles of β , called nucleation, and
- (ii) the increase in size of these stable particles, called growth.

A distinction between the nucleation and the growth stages of a transformation becomes necessary for the following reason. When a β particle is formed, a *new interface* is created between the particle and the liquid. Like all surfaces, this interface has a positive energy, which must be supplied during the transformation process. *A tiny particle has a large surface area to volume ratio* and can therefore be unstable. For a spherical particle of radius r ,

$$\text{Surface area/volume} = (4\pi r^2)/(4\pi r^3/3) = 3/r \quad (9.3)$$

As $r \rightarrow 0$, this ratio can become very large and the energy of the surface can effectively prevent the initial formation of a tiny particle. A particle is said to have nucleated, when it becomes stable and will not disappear due to thermal fluctuations. Once the particle has attained this stage, it can grow further with a continuous decrease in energy. The surface energy is no longer a dominant factor in the growth process.

9.2 The Nucleation Kinetics

In *homogeneous nucleation*, the probability of nucleation occurring at any given site is identical to that at any other site within the volume of the parent phase. We will consider this type of nucleation first. If Δf is the free energy change accompanying the formation of a spherical new phase particle, we can write

$$\Delta f = \frac{4}{3} \pi r^3 \Delta g + 4 \pi r^2 \gamma \quad (9.4)$$

where r is the radius of the particle, Δg is the Gibbs free energy change per unit volume and γ is the surface energy per unit area of the interface separating the parent and the product phases. The surface energy term is always positive. If Δg is negative, the function Δf passes through a maximum. Initially as the new phase particle starts to form, the energy of the system increases, as the surface energy term is dominant. At the maximum, the variations with r of the surface energy and the volume (Gibbs) free energy exactly balance each other. Thereafter, the variation in the volume term becomes dominant and, as this term is negative, there is a continuous decrease in the energy of the system.

By setting $(df/dr) = 0$, the values corresponding to the maximum, called the critical values and denoted by the superscript *, are obtained:

$$r^* = -\frac{2\gamma}{\Delta g} \quad (9.5)$$

$$\Delta f^* = \frac{16}{3} \pi \gamma^3 / (\Delta g)^2 \quad (9.6)$$

Δf is plotted as a function of r for different temperatures in Fig. 9.3. The critical values of the free energy of nucleation Δf^* and of the particle radius r^* are

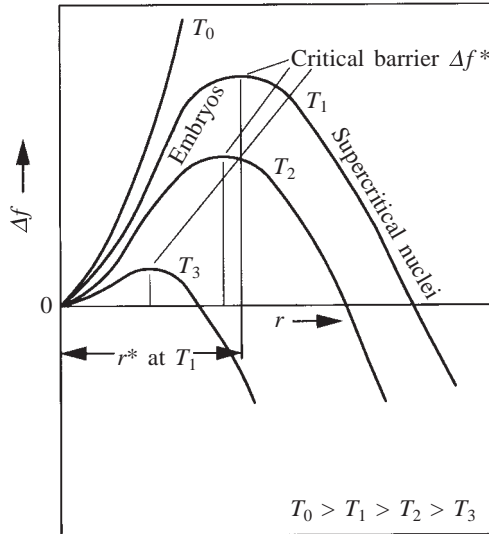


Fig. 9.3 The free energy change Δf during nucleation, as a function of the particle radius r at different temperatures.

indicated in the figure. Particles which are smaller than the critical size are called *embryos*. Those which are larger than the critical size are called *nuclei* (not to be confused with atomic nuclei). The critical sized particle is a *critical nucleus*. As Δg becomes more negative with a lowering of the temperature, the critical condition occurs at smaller values of Δf and r .

For the $L \rightarrow \beta$ transformation, the following simplification usually holds good:

$$\Delta g = \Delta h \frac{T_m - T}{T_m} = \Delta h \frac{\Delta T}{T_m} \quad (9.7)$$

where Δh is the enthalpy change (the heat of reaction) per unit volume of the product assumed to be independent of temperature. T_m is the equilibrium melting point at which $\Delta g = 0$, T is the transformation temperature and ΔT is the degree of supercooling. Combining Eqs. (9.6) and (9.7), we obtain

$$\Delta f^* = \frac{16}{3} \pi \gamma^3 \frac{T_m^2}{(\Delta h)^2 (\Delta T)^2} \quad (9.8)$$

Example 9.1 Calculate the critical free energy of nucleation of ice from water at (i) 0°C , (ii) -5°C , and (iii) -40°C . Also, calculate the critical radius at each temperature. The enthalpy of fusion of ice is 6.02 kJ mol^{-1} . The energy of the ice-water interface, 0.076 J m^{-2} , can be taken to be independent of temperature.

Solution (i) At 0°C , there is no supercooling, i.e., $\Delta T = 0$. So, $\Delta f^* = \infty$. There is no critical radius.

(ii) At -5°C , taking the molar volume of ice as 19 cm^3 ,

$$\begin{aligned} \Delta f^* &= (16 \times 3.142 \times 0.076^3 \times 273^2) / \{3 \times (6020 \times 10^6 / 19)^2 \times 5^2\} \\ &= 2.2 \times 10^{-16} \text{ J} \\ r^* &= (2 \times 0.076 \times 273) / \{-(6020 \times 10^6 / 19) \times 5\} \text{ m} \\ &= 262 \text{ \AA} \end{aligned}$$

(iii) At -40°C , the degree of supercooling is eight times larger than at -5°C .

Therefore, Δf^* is smaller by a factor of 8^2 .

$$\begin{aligned} \Delta f^* &= 3.4 \times 10^{-18} \text{ J} \\ r^* &= 33 \text{ \AA} \end{aligned}$$

The Maxwell-Boltzmann statistics can be used to estimate the number of critical sized particles in the parent phase. Let the total number of particles per unit volume of the parent phase be N_t . Then the number of critical sized particles N^* is given by

$$N^* = N_t \exp \left(\frac{-\Delta f^*}{RT} \right) \quad (9.9)$$

where Δf^* is expressed in units of joule per mole. The process of nucleation is to be identified with that unit step which makes the critical sized particle to become just supercritical.

Let there be s^* atoms in the parent phase facing the critical sized particle across the interface, as shown in Fig. 9.4. If anyone of these jumps from the parent

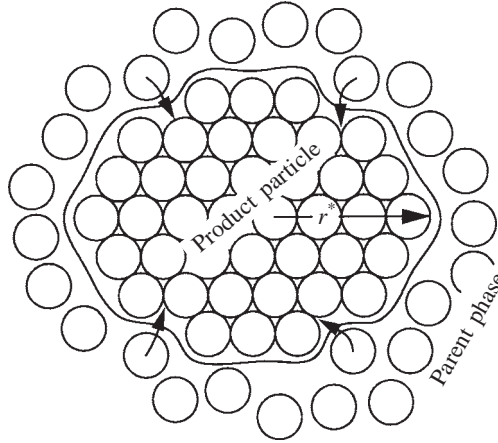


Fig. 9.4 The transfer of atoms from the parent phase to the product particle of critical size r^* across the interface.

phase to the product phase, the particle becomes *just supercritical and is said to have nucleated*. Any further addition of atoms to the particle is to be considered as the subsequent process of growth. The frequency ν' with which anyone of the s^* atoms can cross the interface to join the product particle is given by

$$\nu' = s^* \nu \exp \left(\frac{-\Delta H_d}{RT} \right) \quad (9.10)$$

where ν is the lattice vibration frequency ($\sim 10^{13} \text{ s}^{-1}$) and ΔH_d is the enthalpy of activation for diffusion across the interface. The rate of formation of the new phase particles or the rate of nucleation, dN/dt ($= I$), is approximately the product of the number of critical sized particles and the frequency with which they can become supercritical:

$$I = N^* \nu' = N_t s^* \nu \exp \left(-\frac{\Delta f^* + \Delta H_d}{RT} \right) \quad (9.11)$$

As a first approximation, in homogeneous nucleation, N_t can be taken as the number of atoms in the parent phase per unit volume. Then the pre-exponential term $N_t s^* \nu$ is in the range $10^{42} \text{ m}^{-3} \text{ s}^{-1}$.

Example 9.2 For a measurable (homogeneous) nucleation rate of $10^6 \text{ m}^{-3} \text{ s}^{-1}$, calculate the value of $(\Delta f^* + \Delta H_d)$ at room temperature. By how much Δf^* should be changed to increase the nucleation rate by a factor of 10^4 to $10^{10} \text{ m}^{-3} \text{ s}^{-1}$?

Solution Substituting the appropriate values in Eq. (9.11), the exponent $= (42 - 6) \times 2.303 = 82.9$. At room temperature (300 K), $(\Delta f^* + \Delta H_d)$ comes out as $82.9 \times 8.314 \times 300 = 207 \text{ kJ mol}^{-1}$.

For a nucleation rate equal to $10^{10} \text{ m}^{-3} \text{ s}^{-1}$, we can similarly write $(\Delta f^* + \Delta H_d) = (42 - 10) \times 2.303 \times 8.314 \times 300 = 184 \text{ kJ mol}^{-1}$, a decrease of 23 kJ mol^{-1} . All this decrease is brought about by a decrease in Δf^* , as the enthalpy of activation for diffusion ΔH_d is constant. If the change in Δf^* is brought about by a change in temperature, the corresponding change in the RT term has also to be taken into account in the calculations.

As already noted, Δf^* is a function of temperature. Its value at the equilibrium temperature ($\Delta T = 0$) is infinite. Therefore, the nucleation rate I as given by Eq. (9.11) is zero at the equilibrium temperature. At temperatures below the equilibrium temperature, Δf^* is finite. I increases with decreasing temperature, as Δf^* decreases. At some degree of supercooling, I reaches a maximum value. Thereafter, the decrease in the thermal energy RT dominates the exponential term in Eq. (9.11) and the nucleation rate decreases with decreasing temperature. It becomes zero again at zero kelvin. This temperature dependence of I is shown in Fig. 9.7.

A change in volume accompanies most phase transformations. When the adjoining parent phase is a fluid, volume changes during transformation can be accommodated by flow in the parent phase. If the transformation occurs entirely within a solid, volume changes will introduce elastic strains in the phases. Like surface energy, the elastic strain energy is also positive and tends to inhibit the transformation. The equation for the free energy change Δf during nucleation should be modified, when strain energy is a factor in a transformation. In such cases, the balance between the surface energy and the strain energy usually results in a plate like product particle.

In a transformation where strain energy is a factor, let the product phase be a disc-shaped particle of semi-thickness c and radius r . The volume of this disc can be approximated to $4\pi r^2 c/3$ and the surface area to $2\pi r^2$. A simple form of the strain energy valid for a number of situations is

$$\varepsilon = A(c/r) \quad (9.12)$$

where ε is the strain energy per unit volume and A is the strain energy parameter that includes the elastic constants of the phases. Now, a modified expression for Δf can be written as follows:

$$\Delta f = \frac{4}{3} \pi r^2 c \Delta g + \frac{4}{3} \pi r^2 c A(c/r) + 2\pi r^2 \gamma \quad (9.13)$$

A plot of Δf as a function of r and c is a surface with a saddle point corresponding to Δf^* . The critical values denoted by the superscript $*$ can be obtained by setting the partial derivatives $(\partial \Delta f / \partial r)$ and $(\partial \Delta f / \partial c)$ equal to zero:

$$c^* = -2\gamma/\Delta g \quad (9.14)$$

$$r^* = 4A\gamma/(\Delta g)^2 \quad (9.15)$$

and

$$\Delta f^* = \frac{32}{3} \pi A^2 \gamma^3 / (\Delta g)^4 \quad (9.16)$$

Example 9.3 A new phase forms in the shape of a disc of radius r and semi-thickness c . If the strain energy per unit volume is $A(c/r)$, show that the free energy of the nucleus of a given volume is a minimum, when $c = \sqrt{\gamma r/A}$, where γ is the interfacial energy.

Solution For minimizing the free energy change Δf at constant particle volume V , substitute $c = V/(4\pi r^2/3)$ in Eq. (9.13):

$$\Delta f = V\Delta g + V^2 A / (4\pi r^3/3) + 2\pi r^2 \gamma$$

Setting $d\Delta f/dr = 0$, we have

$$-9V^2 A / (4\pi r^4) = -4\pi r \gamma$$

Resubstituting $V = 4\pi r^2 c/3$, we get

$$c^2 A = \gamma r, \quad c = \sqrt{\gamma r/A}$$

In *heterogeneous nucleation*, the probability of nucleation occurring at certain preferred sites in the assembly is much greater than that at other sites. During solidification of a liquid, inclusions of foreign particles in the liquid and the walls of the container holding the liquid provide preferred nucleation sites. In a solid-solid transformation, foreign inclusions, grain boundaries, interfaces, stacking faults and dislocations can provide sites for preferred nucleation. As an example, consider the nucleation of β from α occurring on the planar surface of a foreign inclusion δ , as illustrated in Fig 9.5.

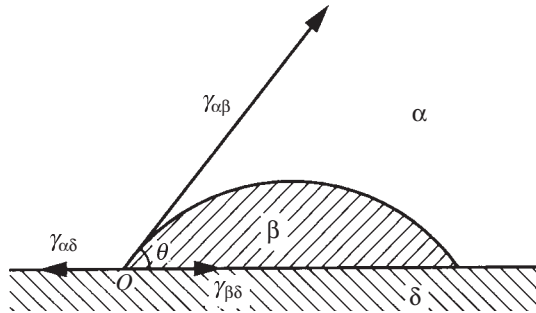


Fig. 9.5 The β particle nucleates on the surface of δ , which is a foreign inclusion in the matrix of α .

The β particle forms in the shape of a half lens on the δ surface. To calculate the nucleation kinetics, we assume that the surface tension forces at point such as O in Fig. 9.5 are in equilibrium:

$$\gamma_{\alpha\delta} = \gamma_{\alpha\beta} \cos \theta + \gamma_{\beta\delta} \quad (9.17)$$

where θ is the contact angle. An expression for Δf can now be written in terms of the volume energy (using the volume of the lens-shaped particle) and the surface energies of the interfaces involved. During nucleation, a new α - β interface of area equal to the curved surface of the lens is created. In addition, a β - δ interface equal in area to the circle of intersection of the lens with the planar δ surface is created. But an α - δ interface of the same area is consumed. This leads to a total of three surface energy terms: two of them positive and one negative. When the expression for Δf is written taking all these terms into account and the equilibrium condition from Eq. (9.17) is substituted, Δf^* comes out to be

$$\Delta f_{\text{het}}^* = \frac{4\pi\gamma_{\alpha\beta}^3}{3(\Delta g)^2} (2 - 3 \cos \theta + \cos^3 \theta) \quad (9.18)$$

Compare Eq. (9.18) with Eq. (9.6) for homogeneous nucleation. The factor outside the brackets on the right side of Eq. (9.18) is $\frac{1}{4} \Delta f_{\text{homo}}^*$. Δf_{het}^* can be derived as a fraction of Δf_{homo}^* for the following cases:

- (i) The product particle makes only *point contact* with the foreign surface. Then clearly, the foreign particle does not play any role in the nucleation process. The contact angle θ for this case is 180° . $\cos 180^\circ = -1$, so, from Eq. (9.18), $\Delta f_{\text{het}}^* = \Delta f_{\text{homo}}^*$.
- (ii) The product particle *completely wets* the foreign surface. It forms a vanishingly thin film on the δ particle. The contact angle $\theta = 0$. $\cos 0^\circ = 1$; so, from Eq. (9.18), $\Delta f_{\text{het}}^* = 0$. There is no barrier to nucleation.
- (iii) The product particle is hemispherical in shape. $\gamma_{\alpha\delta} = \gamma_{\beta\delta}$. Hence, $\theta = 90^\circ$, $\cos \theta = 0$ and $\Delta f_{\text{het}}^* = \frac{1}{2} \Delta f_{\text{homo}}^*$.

This provides us a clue for *selecting a heterogeneous nucleation agent*, when a nucleation process is to be deliberately induced. A small value of the contact angle θ promotes heterogeneous nucleation. For a given value of the energy of the α - β interface, θ can be minimized, by choosing a nucleating agent δ that would form a low energy β - δ interface. If the crystal structure of β and δ are similar and the lattice parameters are nearly equal, $\gamma_{\beta\delta}$ can be kept quite small and this criterion is used in selecting a heterogeneous nucleating agent. For example, the purpose of seeding rain-bearing clouds with AgI or NaCl crystals is to provide for the easy nucleation of the ice crystals, which subsequently melt to form the rain drops. AgI and NaCl have atomic planes that match those of the ice crystal to provide a low energy interface. In a similar way, the nucleation rate of ice crystals can be increased and thereby the size of the product particles can be reduced, to minimize the damage caused by a hailstorm. Nickel is used as a heterogeneous nucleating agent in the production of artificial diamonds from graphite. It is interesting to note that both nickel and

diamond are referred to the FCC space lattice, with a lattice parameter of 3.52 Å and 3.57 Å, respectively.

Example 9.4 Calculate Δf_{het}^* as a fraction of Δf_{homo}^* , when the interfaces shown in Fig. 9.5 have the following energies:

$$\gamma_{\alpha\beta} = 0.5 \text{ J m}^{-2}, \gamma_{\alpha\delta} = 0.5 \text{ J m}^{-2}, \text{ and } \gamma_{\beta\delta} = 0.01 \text{ J m}^{-2}$$

Solution From surface tension equilibrium [Eq. (9.17)],

$$\cos \theta = (\gamma_{\alpha\delta} - \gamma_{\beta\delta})/\gamma_{\alpha\beta} = (0.5 - 0.01)/0.5 = 0.98$$

From Eq. (9.18),

$$\Delta f_{\text{het}}^* = \frac{1}{4} \Delta f_{\text{homo}}^* (2 - 3 \times 0.98 + 0.98^3) = 0.0003 \Delta f_{\text{homo}}^*$$

The nucleation barrier virtually disappears due to the very low energy of the interface between β and δ .

The rate of heterogeneous nucleation can be expressed in a form similar to that of Eq. (9.11). In addition to the difference in the Δf^* term as discussed above, the pre-exponential term will now include the number of preferred nucleation sites only. This is usually many orders of magnitude smaller than the number of potential sites available for nucleation in the homogeneous case. A typical value here for the pre-exponential term is $\sim 10^{26} \text{ m}^{-3} \text{ s}^{-1}$, as compared to $10^{42} \text{ m}^{-3} \text{ s}^{-1}$ for homogeneous nucleation. In spite of this large difference in the pre-exponential term, $I_{\text{het}} > I_{\text{homo}}$, because the result that $\Delta f_{\text{het}}^* < \Delta f_{\text{homo}}^*$ dominates the nucleation kinetics through the exponential term.

9.3 The Growth and the Overall Transformation Kinetics

Growth is the increase in size of the product particle *after* it has nucleated. Growth usually occurs by the thermally activated jump of atoms from the parent phase to the product phase. Consider the transformation $\alpha \rightarrow \beta$. As illustrated in Fig. 9.6, the unit step in the growth process consists of an atom leaving the α

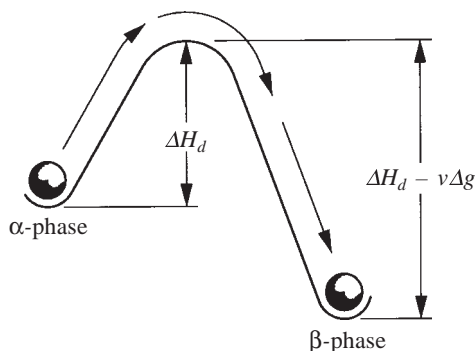


Fig. 9.6 Unit step in the growth of a β -particle into the α -matrix.

matrix and jumping across the α - β interface to join the product phase β . The activation barrier for the $\alpha \rightarrow \beta$ jump is the same as the activation barrier for diffusion across the interface, ΔH_d .

At the equilibrium temperature, α and β phases have the same free energy, so that the frequency of jumps from α to β is equal to that from β to α . The net growth rate is zero. At lower temperatures, the β -phase has a lower free energy and the barrier for the reverse $\beta \rightarrow \alpha$ jump is larger: $\Delta H_d - v\Delta g$, where v is the volume per atom and ΔH_d is in units of J atom⁻¹. Now there is a net flow of atoms from the α -phase to β -phase. The growth rate can be expressed as $U = dr/dt$, where r is the radius of the particle.

As a function of temperature, the growth rate first increases with increasing degree of supercooling, but eventually starts to decrease as the thermal energy RT falls. In Fig. 9.7, the temperature dependence of the growth rate U is shown along with the temperature dependence of the nucleation rate I . As seen from the figure, the growth rate is also a maximum at some intermediate degree of supercooling. The maximum in the growth rate usually occurs at a higher temperature than the maximum in the nucleation rate.

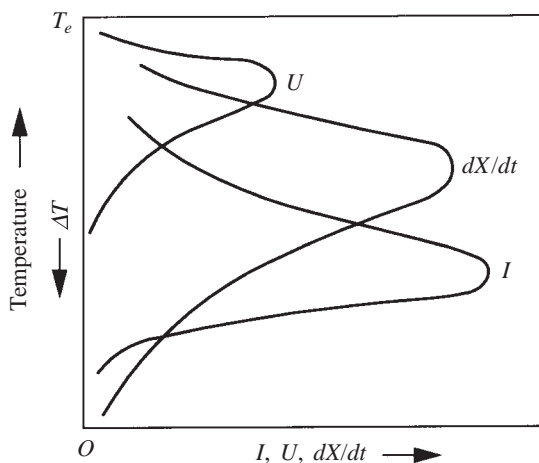


Fig. 9.7 The temperature dependence of the nucleation rate I , the growth rate U , and the transformation rate dX/dt .

The *overall transformation kinetics* can be described as some function of the growth rate and the nucleation rate:

$$dX/dt = f(U, I) \quad (9.19)$$

where dX/dt is the rate at which the fraction X of the new phase increases with time t . Assuming that a constant rate of nucleation occurs randomly in the untransformed phase and that the product particles grow at a constant rate as spheres, till impingement with neighbouring particles occurs, the fraction X can be derived as a function of t :

$$X = 1 - \exp\left(-\frac{\pi}{3} I U^3 t^4\right) \quad (9.20)$$

A plot of X versus t shown in Fig. 9.8 has the sigmoidal form.

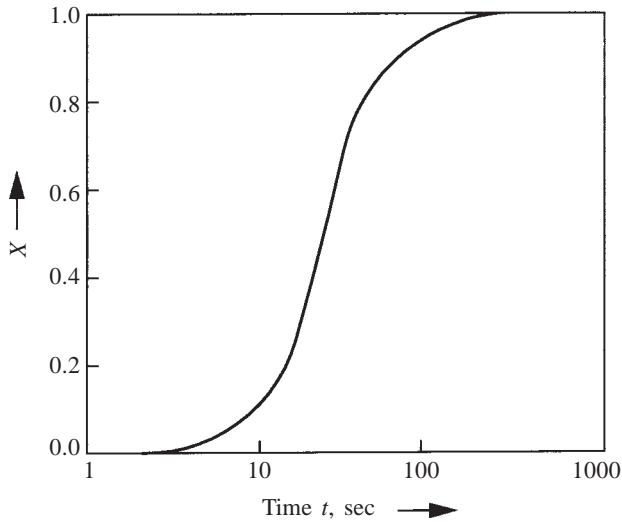


Fig. 9.8 The fraction transformed X is plotted against t . The curve has a sigmoidal form.

As the transformation rate is a function of both the nucleation and the growth rates, the dX/dt curve has the same shape as the U and I curves of Fig. 9.7. It also passes through a maximum at some intermediate degree of supercooling. The temperature of the maximum transformation rate lies between the temperatures of maximum growth rate and maximum nucleation rate, as can be seen from Fig. 9.7. It is often convenient to plot the time taken at various temperatures for a fixed fraction of transformation X , in the form of a T-T-T (*time-temperature-transformation*) curve. The T-T-T curve is inversely related to the dX/dt curve. It has, therefore, a C shape, which is the inverse of the dX/dt curve. The nose of the C curve corresponds to the minimum time for a specified fraction (or per cent) of transformation. The nose temperature is the temperature at which dX/dt is a maximum. A T-T-T diagram for austenite-to-pearlite transformation in a steel is shown in Fig. 9.9 with two C curves for 1% and 99% transformations, respectively.

APPLICATIONS

Some typical applications of the principles of nucleation and growth are discussed below, as examples of how the control of a desired structure can be achieved in a material.

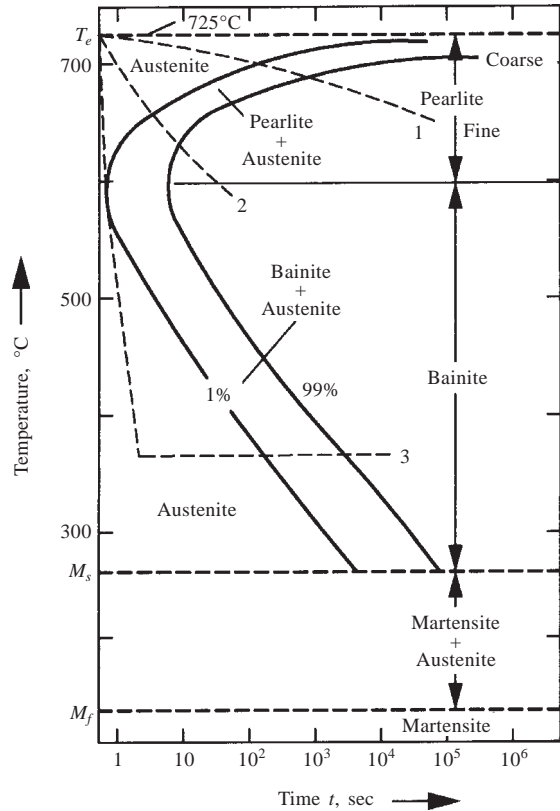
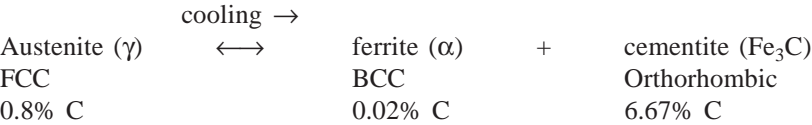


Fig. 9.9 The T-T-T diagram for a eutectoid steel.

9.4 Transformations in Steel

The phase transformations in steel are of great technological importance. The heat treatment of steel to obtain the optimum properties is dependent on the mechanism and the kinetics of phase transformations.

Consider the binary phase diagram of Fe and Fe₃C shown in Fig. 7.11. The following eutectoid reaction occurs on cooling a steel of 0.8% carbon through the eutectoid temperature (725°C):



The transformation takes place with a compositional change. The FCC austenite (γ) of 0.8% carbon decomposes to a mixture of BCC ferrite (α) of 0.02% carbon and orthorhombic cementite (Fe₃C) of 6.67% carbon. This mixture is called *pearlite*. This compositional change is brought about by the *diffusion of interstitial carbon* (which is much faster than the diffusion of iron).

As the product phase particle (ferrite or cementite) grows, carbon diffuses away from or into the transforming region.

The time-temperature-transformation characteristics of the eutectoid steel are shown on a T-T-T diagram in Fig. 9.9. A slow cooling rate from the eutectoid temperature T_e as indicated by curve 1 in Fig. 9.9 yields *coarse pearlite*—a mixture of relatively coarse crystals of ferrite and cementite like that shown in Fig. 7.12. Such a slow cooling is called *annealing*. Here, sufficient time is available for the carbon in the austenite to diffuse and redistribute itself to 0.02% in ferrite and 6.67% in cementite. Small plate-like regions of austenite transform into pairs of parallel, plate-like crystals of ferrite and cementite. The diffusion distance of carbon is approximately half the thickness of a pair of plates. During annealing, the effective transformation temperature is just below T_e , where the rate of growth is rapid compared to the rate of nucleation, recall Fig. 9.7. Such a combination of growth and nucleation rates promotes the formation of coarse crystals.

A somewhat faster cooling rate obtained by *air cooling* lowers the effective range of the transformation temperature. The range is now nearer to the maximum nucleation rate where the growth rate is not so high, see Fig. 9.7. This combination of rates yields *finer crystals* of ferrite and cementite. Curve 2 in Fig. 9.9 corresponds to air cooling. The steel is said to be *normalized* to yield fine pearlite. The diffusion rate of carbon decreases with the decrease in the effective transformation temperature. This effect is, however, compensated by the fact that here the carbon atoms need to diffuse over shorter distances than in coarse pearlite, as the plates are thinner. There is also the additional advantage of a larger undercooling.

The transformation of austenite to pearlite is suppressed by rapid cooling past the nose of the C curve, as represented by cooling curve 3 in Fig. 9.9. If the austenite is then held isothermally below the nose temperature, it transforms to bainite. *Bainite* is also a mixture of ferrite and cementite. This mixture has a very fine distribution of the two phases but has a different morphology as compared to pearlite. Here, the cementite particles are short needles embedded in plates of ferrite. (The nomenclature of the microconstituents in steel is dependent on the shapes of the product phases. Only the lamellar arrangement of pairs of parallel, plate-like crystals of ferrite and cementite is called pearlite.) The diffusion rate of carbon becomes still less below the nose of the C curve. This causes the very fine distribution of the phases in bainite.

The pearlite and the bainite transformations can be represented by separate C curves. In the case of plain carbon steels, the C curves overlap and only the top part of the pearlite C curve and the bottom part of the bainite C curve are evident in Fig. 9.9. Alloy steels contain other substitutional alloying elements in addition to interstitial carbon. The T-T-T diagram for a Ni–Cr–Mo low alloy steel is shown in Fig. 9.10 and has two separate C curves.

If the metastable austenite is cooled rapidly below the bainitic range, the available thermal energy decreases. The diffusion of carbon required to bring about the compositional changes becomes negligible. The austenite has been supercooled by nearly 500°C and there is a large driving force to convert the FCC structure to BCC. Now a new phase transformation that does not require the diffusion of carbon sets in. Austenite transforms to martensite (α').

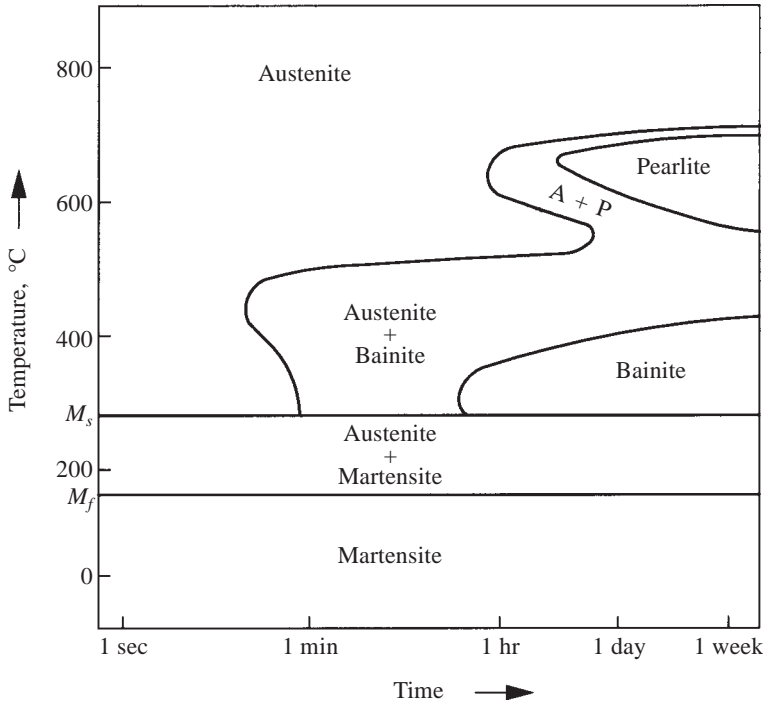
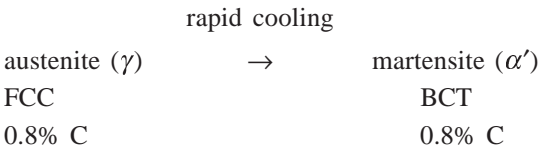


Fig. 9.10 T-T-T diagram for Ni-Cr-Mo low alloy steel depicting two C-curves.



In most steels, the amount of martensite that forms is a function of the temperature to which the austenite is cooled and not a function of time. This is in contrast to the pearlitic and bainitic transformations. The temperature at which martensite starts to form is called M_s (martensite start). The temperature at which the transformation is virtually complete is called M_f (martensite finish). These temperatures are indicated in both Figs. 9.9 and 9.10. Within the martensitic range, the time scale in these figures is not relevant. Figure 9.11 shows plates of martensite in a matrix of austenite.

The martensitic transformation occurs without a compositional change. The transformation is a process of shear, that occurs without any need for diffusion. The atomic movements that bring about the transformation are only a fraction of an interatomic distance. The shear changes the *shape* of the transforming region. This results in a considerable amount of strain energy, due to the resistance of the matrix (parent phase) to the shape change. As a result of the strain energy factor, martensite has a plate-like shape, as seen in Fig. 9.11.

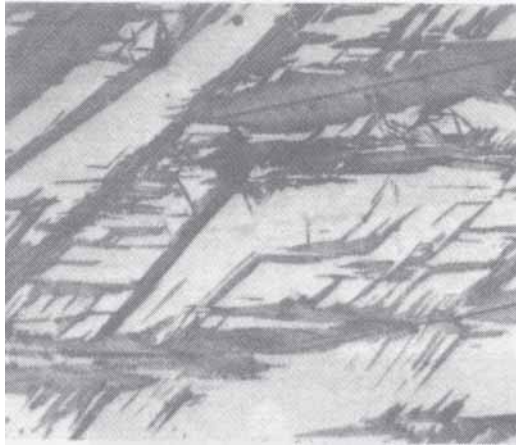


Fig. 9.11 Martensite plates in an austenite matrix at 1000 \times .

Austenite is an interstitial solution of carbon in FCC iron, the carbon atoms occupying the *octahedral voids* of the FCC structure. These voids are located at the body centre and at the twelve edge centres of the FCC unit cell. With the carbon content of the order of 1%, only a small fraction of the total number of octahedral voids is occupied by the carbon atoms. During martensitic transformation, the shear process converts the FCC structure of austenite to the BCT structure of martensite. The mechanism of the transformation is illustrated in Fig. 9.12.

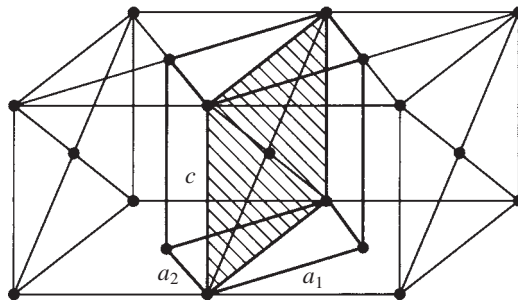


Fig. 9.12 A BCT unit cell outlined in two adjacent FCC unit cells. (The face centred atoms on the front and back faces of the FCC cells are not shown.)

A BCT unit cell is outlined in two adjacent FCC cells. During the transformation, the vertical c -axis of the BCT cell (which is the same as the cube edge) contracts by about 20% and the horizontal a_1 and a_2 axes (which are half of the face diagonal of the cube) expand by about 12% each, so that the BCT cell takes on the cell dimensions characteristic of the product martensitic phase. The iron atoms move only by a fraction of the interatomic distance. There are no individual jumps of atoms from site to site, characteristic of the diffusion process. As the carbon atoms are all situated at the middle of the c -axis (or equivalent positions), the shear displacement of the iron atoms during

contraction of the c -axis is obstructed by the carbon in between. Note that there are no carbon atoms at the middle of the a_1 and a_2 axes. This obstruction along the c -axis results in a tetragonal product, with c/a ratio slightly greater than unity. The c/a ratio of the BCT martensite is a function of the carbon content and varies from 1.0 at 0.0% carbon (BCC martensite) to 1.08 at 1.2% carbon. When the carbon atoms are not present (i.e., in pure iron), on rapid quenching from the FCC region, the shear process converts the FCC to BCC iron. Here, the c -axis contracts and the a_1 and a_2 axes expand to a sufficient degree to make them all equal.

The hardness of martensite is a function of its carbon content. It increases rapidly with increasing carbon content, reaching a more or less constant value of 65 on the R_c scale at about 0.6% carbon, Fig. 9.13. The approximate hardness

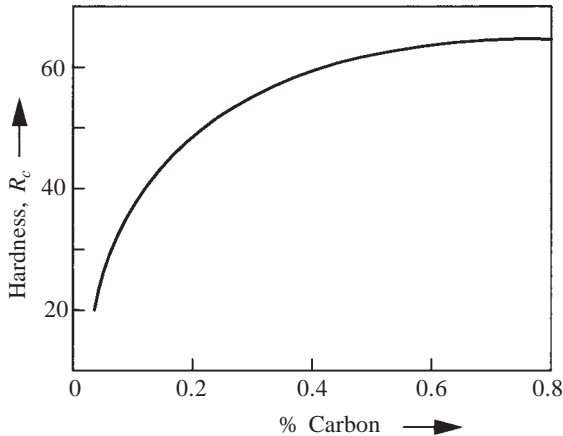


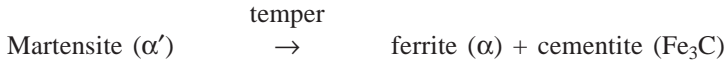
Fig. 9.13 The hardness of martensite on the R_c scale as a function of the carbon content of a steel.

values of the various transformation products in an eutectoid steel are given in Table 9.2. The corresponding tensile strengths are also indicated. *The tensile strength increases with increasing hardness. However, when the hardness is a very high value such as R_c 65, the steel is brittle, resulting in poor tensile strength.*

TABLE 9.2
Properties of the Transformation Products in a 0.8% Carbon Steel

Constituent	Hardness, R_c scale	Tensile strength, MN m^{-2}
Coarse pearlite	15	710
Fine pearlite	30	990
Bainite	45	1470
Martensite	65	—
Martensite tempered at 250°C	55	1990

Martensite is an extremely hard and brittle phase. Its presence is desirable in tool steels but it is too brittle to be useful in the as-quenched condition. It is reheated to reduce its brittleness, without much loss of its hardness. This process is called *tempering*. During tempering, the metastable martensite decomposes into the more stable products of ferrite and cementite by the process of carbon diffusion.



As very fine cementite particles form, the excess carbon in martensite is gradually lost. The decrease in hardness and increase in ductility during tempering can be controlled quite accurately, by a proper choice of tempering temperature and time. The final hardness (and related ductility) of the quenched steel depends on the application. For example, a tool for cutting glass can be used in the as-quenched condition with a hardness of over R_c 65. A chisel or a punch requires some shock resistance in addition to hardness. The tool steel for such applications is quenched and then tempered to yield a final hardness in the range R_c 45–55.

Quenching to obtain martensite induces stresses in the steel. As the heat is extracted by the quenching medium through the surface, there is a temperature gradient from the inside to the outside of the component. This gradient induces not only *thermal stresses* but also *residual stresses* in the steel, as the austenite transformation does not occur simultaneously throughout the cross-section. If the size of the component is large, the stresses may be severe enough to induce cracks in the steel. The presence of cracks or residual stresses is not desirable in most applications.

In a steel containing only carbon, the nose of the C curve is close to the temperature axis (see Fig. 9.9) and a very fast cooling rate (water quenching) is necessary to avoid transformation to pearlite. *Alloying elements* such as Mn, Ni, Cr, V, W and Mo added to steel, have the effect of *shifting the nose of the C curve or C curves to the right*. Compare the nose positions in Figs. 9.9 and 9.10. To obtain martensite, a slower cooling rate (oil quenching) can be adopted for low alloy steels as compared to plain carbon steels. This enables the formation of martensite without cracking the steel. This is one of the main functions of alloying elements in steel. They are said to increase the *hardenability*, which is the ability to harden without having to resort to drastic quenching.

Another function of alloying elements is to provide a fine distribution of alloy carbides during tempering. High speed steel contains very fine tungsten carbide particles after tempering and the tool made out of this steel retains its strength at elevated temperatures. It continues to machine, even when it gets red hot! Cutting speeds can be more than ten times higher for the high speed steel as compared to a carbon steel, resulting in a substantial saving in production time.*

The contributions of alloying elements to the properties of steel are summarized in Table 9.3.

*Still higher cutting speeds are obtained in tungsten carbide tools. These have about 80% of brittle WC particles embedded in a ductile matrix of cobalt.

TABLE 9.3
Contributions of Alloying Elements to the Properties of Steel

Alloying element	Cr	Co	Mn	Mo	Ni	W	V
Hardenability	++	—	++	++	+	++	+++
High temperature strength	+	+	—	++	—	++	++
Ductility and toughness	—	—	+	—	++	—	—
Abrasion resistance	+	—	—	+	—	++	+
Fine austenite grain size	—	—	—	+	—	+	+++
Corrosion resistance	++	—	—	+	+	—	—

Some special heat treatment processes are used to enhance the properties of steel. *Martempering* consists of quenching the steel from the austenitizing temperature to a bath just above M_s , holding in the quenching medium until the temperature of the steel becomes uniform throughout the cross-section and then cooling it in air to produce martensite. The martensite is then tempered as required. As austenite transforms to martensite simultaneously throughout the steel in this process, the distortion in quenching is minimized. This induces greater toughness in the steel. *Austempering* refers to quenching a steel to a temperature below the nose and holding it isothermally to produce *bainite*. The austempered steel possesses greater ductility and toughness than one quenched and tempered in the conventional manner to give the same hardness. *Ausforming* refers to working the steel plastically in the temperature range below the nose and then cooling it to produce martensite. The martensitic plates formed are smaller in size, yielding greater ductility for the same strength after the tempering treatment. An ausformed low alloy steel may have a tensile strength as high as 3000 MN m^{-2} , with an adequate ductility of 10% elongation in a tensile test.

9.5 Precipitation Processes

Solubility in the solid state often decreases with decreasing temperature. A solid solution that is stable at higher temperatures can become unstable and supersaturated with respect to the solute at lower temperatures. Then the free energy change favours the rejection of the excess solute in the form of precipitate particles, the matrix solution becoming poorer in solute and approaching equilibrium solubility characteristic of the lower temperature. The precipitation reaction can be written in the following form:



The precipitate particle nucleates and grows by the diffusion of solute atoms into it from the matrix phase.

Precipitation in metallic systems can be illustrated with the transformation in *duralumin*. Duralumin is an alloy of aluminium with about 4% copper and smaller quantities of other alloying elements. Figure 9.14 shows the aluminium rich end of the Al–Cu phase diagram. The alloy with 4% Cu exists as a single phase α solid solution at 550°C and at room temperature as a mixture of α (with less than 0.5% Cu in solution) and an intermetallic compound, CuAl_2 (θ) with

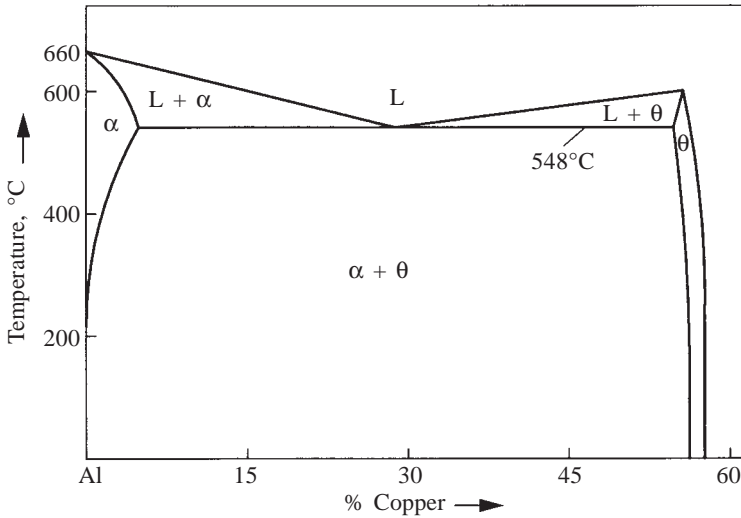
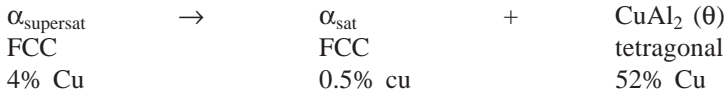


Fig. 9.14 Aluminium rich end of the Al-Cu binary phase diagram.

52% Cu. On slow cooling from 550°C to room temperature, the α phase becomes supersaturated with respect to the copper solute and rejects the excess copper as precipitate particles of θ .



These particles are relatively coarse in size and can be identified under the optical microscope.

On rapid cooling of the alloy, there is not enough time for diffusion of copper atoms to form the precipitate particles. Therefore, a supersaturated solid solution is obtained at room temperature. As a function of time at room temperature and at higher temperatures up to 200°C, the diffusion of copper may take place and the precipitate particles can form. Due to the limited diffusion that is possible in this temperature range, *very fine* precipitate particles are obtained, with the diffusion distance in the range of 100–500 Å. Such precipitate particles are not visible under the optical microscope due to their fine size. *The finely distributed particles effectively hinder the motion of dislocations in the aluminium matrix and increase the strength of the alloy*, as discussed in greater detail in Chap. 11. Controlling the progress of the precipitation reaction so as to obtain the maximum hindrance to dislocation motion and hence the maximum strength of the alloy is the aim of the heat treatment of this alloy. The variation in the hardness of the alloy at three temperatures of ageing between 20°C and 200°C is shown in Fig 9.15. Hardness is a measure of the strength of the alloy. Obviously, the maximum hardness and strength is obtained only for the right combination of temperature and time. This combination is a function of the copper content and the other impurities in the aluminium. Note that the ageing

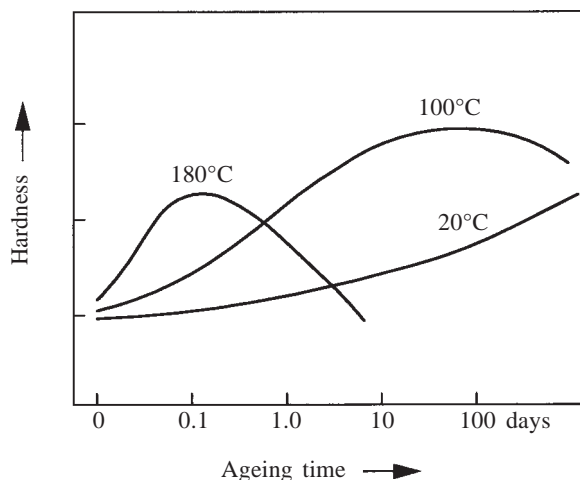


Fig. 9.15 Ageing curves for a duralumin alloy.

time axis in Fig. 9.15 is on *log scale*, so that the optimum peak is not obtained at room temperature, even after several years!

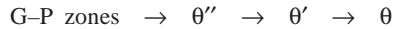
The increase in the hardness in the initial stages of the ageing curves can be attributed to the precipitation process taking place progressively. After reaching a peak value, the hardness starts to decrease. This phenomenon is called *overageing*. As the precipitate particles are very fine in size, they have a high surface to volume ratio. Therefore, they have a tendency to coalesce or *coarsen*, that is, a number of small particles merge to form a large particle. For a given volume fraction of precipitate particles, the coarsening process decreases the total number of precipitate particles and increases the interparticle spacing. The hindrance to the dislocation motion is thereby reduced, accounting for the decrease in hardness beyond the peak in the curve. A temperature between 100 and 180°C (see Fig. 9.15) would be the optimum ageing temperature, where the ageing time is not unduly large. An alloy aged to the optimum peak at such a temperature would not overage during service at room temperature.

Duralumin alloy corresponding to U.S. specification 2024-T6 can have a tensile strength of 500 MN m^{-2} ($\sim 50 \text{ kgf/mm}^2$), as compared to the strength of aluminium, which is about 100 MN m^{-2} . Another age-hardening alloy of Al–Zn of specification 7075-T6 has a tensile strength of 550 MN m^{-2} . These two alloys are technologically important, as they are used as aircraft structural material. An increase in strength by a factor of five as in the above cases makes a vital difference to the economics of air transportation, where the ratio of the dead load (nonpaying load) to the pay load is an important criterion in the choice of a material of construction. The density of these alloys is not very much higher than that of aluminium, as the alloy content is less than 10%.

In recent years, lithium, the density of which is only 20% of that of aluminium, has been added as an alloying element to precipitation-hardened aluminium alloys. In addition to lowering the density, lithium increases the

elastic modulus of aluminium (an anomalous effect!). When Al–Li alloys are fully developed, it is expected that the aircraft structural material in many military aircrafts will be composed of Al–Li alloys.

In a precipitation process, if the crystal structures of the precipitate and the matrix phases are such that a good match can be obtained at the interface, the nucleation barrier is reduced considerably by the formation of a low energy interface between the precipitate and the matrix. If the crystallographic matching is poor, a metastable precipitate of a crystal structure different from the stable one, but which matches better with the parent crystal at the interface, may form due to favourable nucleation conditions. In fact, in the duralumin alloy, the precipitation process is characterized by a series of transition precipitates. They form at different ageing temperatures evolving through the following stages:



G–P zones are small regions in the parent matrix, where the solute atom (copper) has segregated. θ'' and θ' are metastable transition precipitates with distinct crystal structures of their own. They both form on $\{100\}$ planes of the parent FCC matrix, with full or partial matching at the interface. θ is the equilibrium stable precipitate of CuAl_2 .

Example 9.5 Aluminium alloy rivets are often age hardenable. During the riveting operation, they must be in the soft condition so that they can be deformed easily. They age harden slowly at ambient temperature. What precautions are necessary to prevent age hardening before the riveting operation?

Solution If the alloy is already in the aged condition, the precipitated particles should first be put into solution by heating to about 550°C . Then the alloy should be quenched to obtain the supersaturated state. The riveting operation must be undertaken immediately. If the operation is to be delayed, then the quenched alloy should be stored under refrigeration to prevent ageing, which will make the rivets hard and difficult to cold form. Lower temperatures slow down the diffusion process, which is necessary for precipitation, and hence prevent the increase in hardness.

Silicate glasses dissolve almost all the chemical elements in the molten state, but on cooling some phases may precipitate due to supersaturation. These precipitate particles are usually fine crystals of a different composition and refractive index from the matrix. By controlling the number, size, distribution and chemical composition of the precipitating phases, *opaque glasses* or glasses with *different colours* can be produced. The scattering of a part of the visible spectrum occurs when *the spacing* between the precipitate particles *matches the wavelength* of that part of the radiation. Colour filters, railroad signals and neon lamp tubes are examples of materials produced by this type of controlled precipitation. TiO_2 precipitated from the melt is an excellent opacifier, as its refractive index is very high compared to that of silicate glass. For high translucency, the second phase particles must have a refractive index not very different from the matrix. NaF and CaF_2 have refractive indices of 1.3 and 1.4 as compared to 1.5 of a silicate glass and hence are used for making *translucent* glass.

9.6 Solidification and Crystallization

Solidification is the transformation of materials from the liquid to the solid crystalline state on cooling. During solidification, the disordered structure of the liquid transforms to the orderly arrangement characteristic of the crystal. When the disordered structure is frozen as it is, the change is called a *glass transition* which is described in the next section. Crystallization is a term used to refer to the transformation to the crystalline state starting from either a liquid or a glass.

It is in general difficult to prevent the crystallization of liquids which have a large value of the heat of fusion Δh . Recall that the activation barrier for nucleation decreases inversely as the square of the heat of fusion, Eq. (9.8). Crystallization is facilitated by a low value of the activation enthalpy ΔH_d for atom transfer across the interface, recall Eq. (9.11). ΔH_d is approximately proportional to the logarithm of the viscosity of the liquid. A low ΔH_d is characteristic of a low viscosity liquid.

The heat of fusion of metals is relatively large, in the range of 10–15 kJ mol⁻¹. The viscosity of liquid metals is relatively low, in the range 0.1–1 Pa s (1–10 poises). Both of these factors favour the rapid crystallization of liquid metals on cooling. In fact, by ordinary methods of quenching, the crystallization of metals cannot be suppressed. *Splat cooling* involving cooling rates as high as 10⁷ K s⁻¹ prevents crystallization in some alloys.

A *fine grain size* in the solidified product is often desirable for obtaining *superior mechanical properties*. As we have already noted, the combination of a high nucleation rate with a slow growth rate yields fine crystals in the product. Increasing the cooling rate increases the nucleation rate, thereby yielding fine crystals. The full potential of this principle is being realized in recent years with techniques such as *rapid solidification processing*, which has resulted in ultrafine grain sizes. The cooling rates employed here (10⁴–10⁵ °C s⁻¹) fall between those in conventional quenching and splat cooling. Grain refinement can also be achieved by adding nucleating agents, which increase the heterogeneous nucleation rate. External nucleating agents offer the only means of grain refinement in cases, where the metal part cannot be cold worked or recrystallized (Sec. 9.8) or where there are no solid state transformations to manipulate.

As each nucleus grows into a crystal, for obtaining a *single crystal* from a melt, *only one nucleus* must be allowed to grow. Single crystals are grown by seeding the melt with a solid piece of crystal of the same material. The seed provides a ready made stable nucleus. It is made to just touch the liquid and is then pulled out slowly. The assembly is kept under a temperature gradient so as to keep the melt above the freezing temperature and the pulled out crystal below the freezing temperature. This avoids any nucleation in the bulk of the melt and also prevents the remelting of the growing crystal. As the seed is pulled out, the new crystal grows in the same orientation as the seed crystal without the need for an interface (grain boundary) between them. The pulling rate is in the range of 0.01 mm s⁻¹. The slow pulling rate promotes the formation of a crystal with very few line imperfections. Such nearly perfect crystals are a prerequisite to ensure proper operation of solid state devices fabricated out of them.

Silicates, borates and phosphates tend to form glasses. As a consequence of the high cation-cation repulsion, they have open structures. For example, in silicates, the tetrahedra are joined at the corners with a certain amount of freedom of position for the tetrahedral units surrounding a central unit. The difference in the total bond energy between a regular and an irregular array of basic units in space is small, as the bond energy is primarily determined by the first neighbours of the central cation within a unit. Therefore, the heat of fusion Δh , which is a measure of the bond energy between the regular and the irregular arrays, is small for open structures. Also, the viscosity (which is related to the activation enthalpy for atom transfer ΔH_d) is high, of the order of 1000 Pa s at the equilibrium freezing temperature. Both these factors are unfavourable for rapid nucleation and growth. The nucleation and the growth rates of crystals are so negligible that, for all practical purposes, there is no crystallization. Addition of certain oxides to silica promotes closer packing and hence increases the tendency to crystallize. As *forming operations are readily carried out in the glassy state*, the silicate glass manufacturer carefully controls the composition of his working material so as to avoid crystallization.

Formed glass articles can be crystallized by a special subsequent heat treatment process to yield better mechanical and thermal properties. In the production of crystallized glass known as *pyroceram* (also called glass-ceramic), heterogeneous nucleating agents such as TiO_2 are first dissolved in the molten silicate. After shaping the material in the glassy state, the TiO_2 is allowed to precipitate as very fine particles. The surfaces of these particles provide easy nucleation sites for crystals of the matrix to form. A large number of crystals are nucleated by keeping the material for a period at the temperature corresponding to the maximum nucleation rate. The material is then heated to the temperature of the maximum growth rate, where the already formed stable nuclei grow to larger sizes. Even at the end of the heat treatment, the crystallization is incomplete and the fine crystals remain embedded in the glassy matrix. This time saving cycle is illustrated in Fig. 9.16. The size of the crystals is about $0.1 \mu\text{m}$, as compared to a typical value for a fine grain size in metals, which is

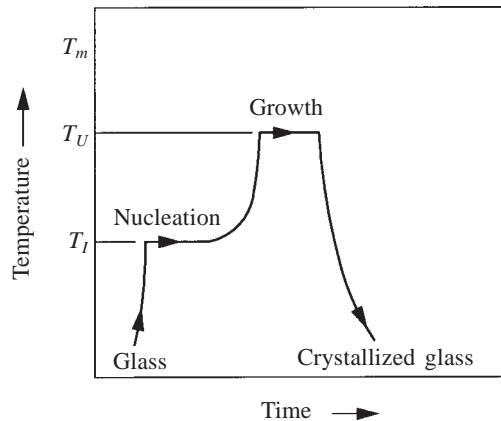


Fig. 9.16 Heat treatment cycle for the crystallization of pyroceram.

10 μm . This *ultrafine grain size* is the reason for the good mechanical and thermal shock resistance of pyroceram. Cookware made of pyroceram can be heated under direct fire on a gas range, without the risk of cracking.

Under industrial casting conditions, the kinetics of solidification is often controlled by heat flow. When a melt is poured into a container or mould, which is at a much lower temperature than the melt, large supercooling occurs in regions where the melt comes into direct contact with the cold container walls. This large supercooling results in a rapid rate of nucleation and a *layer of fine crystals* is formed adjacent to the wall. The latent heat released in the formation of this layer increases the temperature and reduces the nucleation rate in the liquid next to it. Heat flows out of the mould through the walls. As a result, the crystals adjacent to the layer of fine crystals at the mould wall grow inwards towards the centre (in a direction opposite to that of the heat flow) producing long columnar crystals. These crystals ultimately meet near the centre of the mould and a cross-section after solidification shows crystals radially branching out from the centre.

Quite often, crystals grow preferentially along certain crystallographic directions, along which the thermal conductivity (and hence the heat extraction) is a maximum. This fact is exploited in certain applications to produce a polycrystalline material with crystals aligned in a particular crystal direction. In magnetic applications, for example, the alignment of crystals enhances the magnetic properties.

9.7 The Glass Transition

Some liquids do not crystallize easily and can be supercooled to become a rigid noncrystalline solid. The changes in viscosity of a soda-lime-silicate on cooling from the liquid range are shown schematically in Fig. 9.17. If the liquid were to

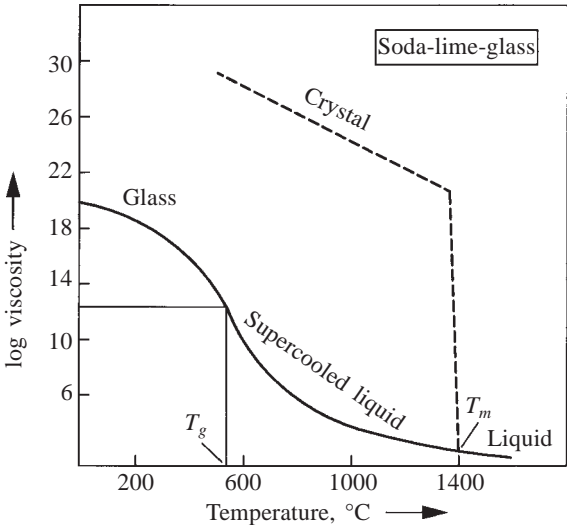


Fig. 9.17 Viscosity changes during cooling of a liquid. Crystallization increases the viscosity abruptly in contrast to the gradual increase that occurs when there is no crystallization.

crystallize, the viscosity would change *abruptly* at the freezing temperature from a low value (about 100 Pa s) in the liquid state to a very high value (about 10^{20} Pa s) in the crystalline state, a change of 18 orders of magnitude at the freezing temperature. However, crystallization does not occur due to kinetic barriers. The viscosity gradually increases with decreasing temperature and attains high values only at low temperatures. The highest rate of change of viscosity with temperature occurs around 10^{12} Pa s at the *glass transition temperature* T_g , see Fig. 9.17. This provides a convenient point for making a distinction between a glass (noncrystalline solid) and a supercooled liquid. In fact, *a solid can be defined as a material with viscosity greater than 10^{12} Pa s, irrespective of whether it is crystalline or not.*

The supercooled liquid, even though it has a higher free energy than the crystals, is in internal equilibrium within itself. That is, the atomic configurations in the supercooled liquid are such that its free energy is a minimum for the liquid-like structure at that temperature. At the glass transition temperature, the thermal energy becomes insufficient for any further configurational adjustments to take place within a reasonable amount of time. The freezing of the first few configurations on approaching T_g favours the freezing of other neighbouring configurations in a cooperative fashion. The range of temperature over which the entire atomic configuration is rendered immobile (except for atomic vibrations about their mean positions characteristic of a solid) is some 10°C for organic polymeric glasses and somewhat larger for inorganic glasses such as silicates and borates.

In glass transitions, the *cooling rate* around T_g is important in the control of properties. For a slower cooling rate, the transition temperature is lower and the specific volume (volume per unit mass) at any temperature below the transition region is also smaller, as illustrated in Fig. 9.18. Obtaining a specific volume as close as possible to the equilibrium volume (which of course is the specific

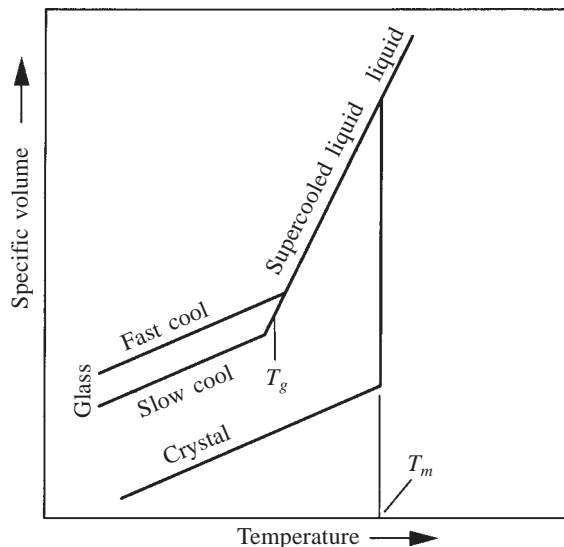


Fig. 9.18 Specific volume of a material as a function of cooling rate.

volume of the crystalline state) is important in all glass applications. Even small amounts of shrinkage over a period of time at service temperatures (which are below T_g) in measuring apparatus such as glass thermometers can lead to important errors. Such shrinkage in optical glasses can lead to nonuniformity in the refractive index and consequent errors in observations made with their aid. For minimum shrinkage in service, the glass should be cooled *as slowly as possible* through the glass transition temperature. A rapidly cooled glass can be reheated and annealed at a temperature just below T_g and cooled slowly, to protect against future shrinkage.

In organic polymers, the glass transition temperature is dependent on the molecular weight of the polymer. Smaller molecular weights correspond to shorter chain lengths and a lower glass transition temperature. Plasticizers increase the distance of separation between chains and reduce the interfering effect on each other and this lowers T_g . PVC cable sheaths, rain coats, etc., are manufactured at room temperature, after lowering the glass transition temperature of polyvinyl chloride by the addition of a plasticizer.

Typical glass transition temperatures T_g of some common long chain polymers are compared with their melting points T_m as illustrated in Table 9.4. The ratio of T_g/T_m lies in the range $\sim 0.4\text{--}0.75$.

TABLE 9.4
Glass Transition Temperatures (T_g) and Melting Points (T_m) for Some Common Long Chain Polymers

Polymer	T_m (K)	T_g (K)	T_g/T_m
LD polyethylene	388	153	0.39
HD polyethylene	410	153	0.37
Polyvinylchloride	465	360	0.37
Polypropylene	445	257	0.58
Polystyrene	513	378	0.74
Polyacrylonitrile	593	380	0.64
6,6 Nylon	538	323	0.60
Polyester	528	348	0.66
Polyisoprene	303	220	0.73

9.8 Recovery, Recrystallization and Grain Growth

Recovery, recrystallization and grain growth are phenomena intimately associated with the annealing of a plastically deformed crystalline material. In crystalline materials, the density of point imperfections and dislocations increases with increasing amount of plastic deformation carried out at temperatures below the range $0.3\text{--}0.5T_m$, where T_m is the melting point in kelvin. Plastic working below $0.3\text{--}0.5T_m$ is called *cold work*. Recall from Chap. 6 that point imperfections and dislocations have strain energy associated with them. Between 1 and 10% of the energy of plastic deformation is stored in

the material in the form of this strain energy. On annealing, that is, on heating the deformed material to higher temperatures and holding, thermal energy comes into play. The material tends to lose the extra strain energy and revert to the original condition before deformation, by the processes of recovery and recrystallization.

During *recovery*, which takes place at low temperatures of annealing, the excess point imperfections that are created during plastic deformation are absorbed at the surface or the grain boundaries or at dislocations by the climbing up process, see Sec. 6.3. Also, random dislocations of opposite sign come together and mutually annihilate each other. Dislocations of the same sign arrange themselves into lower energy configurations, such as tilt and twist boundaries. However, the decrease in the dislocation density during recovery is *not* substantial.

Recrystallization is the process of nucleation and growth of new, strain-free crystals, which replace all the deformed crystals of the worked material. It starts on heating to temperatures in the range of $0.3\text{--}0.5T_m$, which is above the recovery range. There is no crystal structure change during recrystallization. As such, recrystallization is not a phase transformation in a strict sense. The free energy change during recrystallization arises from the excess strain energy of the deformed material as compared to the undeformed material.

Example 9.6 The dislocation density in a copper sample is increased by cold working from 10^9 m m^{-3} to 10^{13} m m^{-3} . Calculate the free energy change during recrystallization.

Solution Strain energy of dislocations in the cold worked copper

$$\begin{aligned} &= \frac{1}{2} \mu b^2 \times 10^{13} \\ &= \frac{1}{2} \times 45.5 \times 10^9 \times 2.55^2 \times 10^{-20} \times 10^{13} \\ &= 14\,800 \text{ J m}^{-3} \end{aligned}$$

We can neglect the strain energy in copper before cold working, as it is four orders of magnitude smaller. The free energy change during recrystallization is then given by

$$\Delta g = -14\,800 \text{ J m}^{-3}$$

The strain energy difference between the cold-worked and the strain-free material is known as the driving force for recrystallization. Nucleation in the usual sense may not occur in recrystallization. An existing grain boundary with local differences in dislocation density on either side may simply migrate into the region of higher dislocation density. The recrystallization temperature is arbitrarily defined as that temperature at which 50% of the material recrystallizes in 1 hr.

Some well-known empirical laws of recrystallization are:

1. The higher is the degree of deformation, the lower is the recrystallization temperature.
2. The finer is the initial grain size, the lower is the recrystallization temperature.
3. Increasing the amount of cold work and decreasing the initial grain size produce finer recrystallized grains.
4. The higher is the temperature of cold working, the less is the strain energy stored in the material. The recrystallization temperature is correspondingly higher.
5. The recrystallization rate increases exponentially with temperature.

The recrystallization temperature is strongly dependent on *the purity of a material*. Very pure materials may recrystallize around $0.3T_m$, while impure materials may recrystallize around $0.5\text{--}0.6T_m$. For example, aluminium of 99.999% purity recrystallizes at 75°C ($348\text{ K} = 0.37T_m$). Commercial aluminium recrystallizes at 275°C ($548\text{ K} = 0.59T_m$). The recrystallization temperature T_r of some pure metals are compared with the melting point T_m as shown in Table 9.5. The ratio of T_r/T_m lies in the range $0.35\text{--}0.5$.

TABLE 9.5
Recrystallization Temperatures (T_r) and Melting Points (T_m)

Metal	T_m (K)	T_r (K)	T_r/T_m
Mg	923	473	0.51
Al	933	423	0.45
Ag	1235	473	0.38
Au	1337	473	0.35
Cu	1358	473	0.35
Ni	1726	873	0.51
Fe	1811	723	0.40
Pt	2042	723	0.35
Mo	2883	1173	0.41
Ta	3269	1273	0.39
W	3683	1473	0.40

During recrystallization, the impurity atoms segregated at the grain boundaries retard their motion and obstruct the processes of nucleation and growth. This *solute drag effect* can be exploited in raising the recrystallization temperature in applications where the increased strength of a cold worked material is to be maintained at the service temperature without letting it to recrystallize.

Recrystallization is also slowed down in the presence of second phase particles. When the particle lies in the migrating boundary during recrystallization, the grain boundary area is less by an amount equal to the cross-sectional area of the particle. When the boundary moves out, it has to pull away from the particle and thereby create new boundary area equal to the cross-section of the particle. This increase in energy manifests itself as a *pinning action of the particle* on the boundary. Consequently, the rate of recrystallization decreases.

Grain growth refers to the increase in the average grain size on further annealing, after all the cold worked material has recrystallized. As a reduction in the grain boundary area per unit volume of the material occurs during grain growth, there is a decrease in the free energy of the material. Consider a curved grain boundary. The atoms on one side of the boundary have on an average more nearest neighbours than on the other side. Therefore, the atoms tend to jump across the boundary to increase their overall bond energy. It is easy to see that the boundary must move towards its centre of curvature for the atoms to go into a position of greater binding. This results in a tendency for larger grains to grow at the expense of smaller grains. As the grains grow larger, the curvature of the boundaries becomes less. The rate of grain growth decreases correspondingly. The state of binding on either side of a planar boundary is the same and, therefore, a planar boundary tends to remain stationary.

In practical applications, grain growth is usually not desirable. Incorporation of impurity atoms (which give rise to the solute drag effect) and insoluble second phase particles (which produce the pinning action on migrating boundaries) are effective in retarding grain growth as well.

The effect on mechanical and some physical properties of the phenomena discussed in this section are summarized in Fig. 9.19. With increasing cold work, the tensile strength increases, but the electrical conductivity and the ductility decrease. On recovery, the electrical conductivity is mostly restored, as it depends mainly on the presence of point imperfections. On recrystallization, the tensile strength decreases and the ductility increases to the values prior to cold working. The microstructural changes are also sketched in Fig. 9.19. During cold work, the grains become elongated in the direction of working. During recrystallization, new equiaxed grains form. During grain growth, these new grains increase in size, but decrease in number.

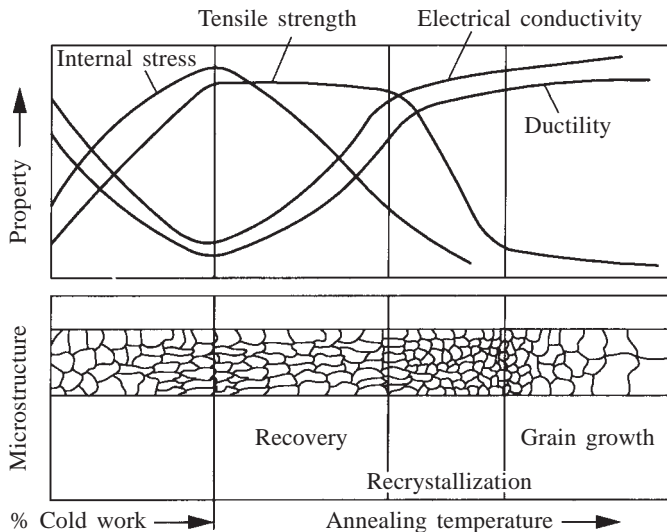


Fig. 9.19 Effect of cold work, recovery, recrystallization and grain growth on some properties of crystalline materials.

SUMMARY

1. Our ability to suppress a phase transformation depends on the time taken by the transformation to go to completion.
2. Phase transformations commonly occur by the processes of nucleation and growth.
3. Even when the volume free energy change is negative, the energy of the interface created between the parent and the product phases inhibits the formation of the product phase in the very early stages, when the surface to volume ratio is high.
4. Strains due to volume and shape changes in solid-solid transformations add to the inhibiting effect of the surface energy.
5. Surface energy requirements can be effectively lowered by choosing a proper nucleating agent to promote heterogeneous nucleation.
6. Most phase transformations that occur on cooling pass through a critical temperature of maximum transformation rate. If the cooling rate is fast enough to avoid transformation in the critical range, the parent phase can be obtained in a highly metastable condition.
7. Our ability to harden steel by the formation of martensite depends on avoiding the pearlitic and bainitic transformations. A good hardenability (ability to harden) is a practical necessity to avoid cracking due to drastic cooling.
8. Proper control of the size and spacing of the precipitate particles that form from the supersaturated solution yields the optimum strength after ageing.
9. Inducing crystallization or preventing it in a glass forming material would depend on the requirements of processing or final properties of the material.
10. A cold worked material undergoes recovery, recrystallization and grain growth at successively higher temperatures of annealing. The retardation of the grain boundary motion is the key factor in raising the recrystallization temperature and preventing grain growth.

PROBLEMS

- 9.1 Derive the results given in Eqs. (9.5)–(9.7).
- 9.2 Calculate the ratio of the surface energy term to the volume energy term in the nucleation energy equation at the critical condition.
Answer: $-3/2$.

- 9.3** Calculate the undercooling that is required for liquid to crystal transformation in tin. The enthalpy of fusion for tin is 0.42 GJ m^{-3} (14 cal/gm). Appreciable nucleation occurs when the free energy of the critical nucleus is $1.5 \times 10^{-19} \text{ J}$. The liquid-crystal interfacial energy is 0.055 J m^{-2} .

Answer: 163 K.

- 9.4** Calculate the critical radius of the copper nucleus during solidification of liquid copper at 983°C . The enthalpy of fusion of copper is 1.88 GJ m^{-3} (50 cal/gm). The liquid-crystal interfacial energy is 0.144 J m^{-2} .

Answer: 21 Å.

- 9.5** In a phase transformation, an appreciable rate of nucleation of $10^6 \text{ m}^{-3} \text{ s}^{-1}$ occurs at room temperature when the critical nucleation energy is $2.07 \times 10^{-19} \text{ J}$. The energy of the interface between the product and parent phases is 0.05 J m^{-2} . Calculate the nucleation rate, if the interfacial energy had been 10% larger than this.

Answer: $6.5 \times 10^{-2} \text{ m}^{-3} \text{ s}^{-1}$.

- 9.6** Derive the results given in Eqs. (9.14)–(9.16).

- 9.7** For a solid-solid phase transformation, the strain energy due to volume change during the transformation is 0.1 GJ m^{-3} . At what magnitude of the free energy change, Δg , the nucleation becomes just feasible?

Answer: 0.1 GJ m^{-3} .

- 9.8** Liquid nucleus forming on a solid surface of the same composition completely wets it. Explain why melting occurs without superheating.

- 9.9** What are the crystal planes that are likely to match atom by atom at the interface in the FCC→HCP transformation? What is the relationship between the lattice parameters of the two crystals for an ideal matching?

- 9.10** Arrange the following carbides in order of their effectiveness as heterogeneous nucleating agents for the crystallization of aluminium. The % misfits for the best matching crystal planes at the interface are given in brackets: TiC (6%), ZrC (14.5%), VC (1.4%), and NbC (8.6%). Give a reason for your sequence.

- 9.11** Explain why the matching of planes at the interface must be only between the product and the nucleating agent and not between the parent phase and the nucleating agent.

- 9.12** The interlamellar spacing of pearlite λ is given by $\lambda = 25/\Delta T \text{ } \mu\text{m}$, where ΔT is the degree of supercooling below the eutectoid temperature. Estimate the temperature at which the pearlite in Fig. 7.12 formed.

Answer: $\sim 700^\circ\text{C}$.

- 9.13** Draw a T-T-T diagram for a 0.6% carbon steel and show by neat sketches the microstructures you would expect for different rates of cooling.

- 9.14** State the simplest heat treatment necessary for steel to convert a 0.8% carbon steel from one structure to another:
- (i) martensite to pearlite,
 - (ii) pearlite to bainite, and
 - (iii) austenite to bainite.
- 9.15** Describe the special features of the martensitic transformation as compared to the other transformations in steels.
- 9.16** Pearlite nucleates heterogeneously at the austenite grain boundaries. Show two C curves for 1% transformation to pearlite—one for a coarse grain size and the other for a fine grain size of austenite.
- 9.17** Referring to the Cu–Zn phase diagram in Fig. 7.10, explain whether it is possible to have a precipitation reaction by quenching a 30% Zn alloy from the α region.
- 9.18** Is an appreciable rate of age hardening of a commercial aluminium alloy always desirable at room temperature? Explain.
- 9.19** Nylon crystallizes at slow rates of cooling but forms glass on quenching in water. Copper always crystallizes even during rapid quenching. Soda-lime-silica forms glass, even when cooled very slowly. Show these differences schematically by drawing C curves for crystallization in these materials.
- 9.20** A polymer that is in the rubbery state has a T_g below room temperature. Explain why this is so.
- 9.21** Some metals can be worked almost indefinitely at room temperature without any noticeable hardening, while others harden at room temperature. Explain the reason for this difference in behaviour, by listing the room temperature in each case as a fraction of T_m for Cu, Ni, Fe, Pb, Sn and Zn.
- 9.22** Tungsten can be cold rolled at 1200°C. Justify this statement.
- 9.23** Hard rolled copper shows 50% recrystallization after heating for 9 min at 135°C or 240 min at 88°C. What is the time required for 50% recrystallization at 50°C?
- Answer:* 6864 min.
- 9.24** Compare the methods available for refining the grain size of copper and mild steel.
- 9.25** Show that the temperature of the maximum rate of nucleation T_{\max} satisfies the following relationship: $d(\Delta f^*)/dT = (\Delta f^* + \Delta H_d)/T_{\max}$, where the symbols have the same meaning as in the text. Suggest a graphical method of determining T_{\max} .
- 9.26** Distinguish between hardness and hardenability of a steel, outlining the factors that influence each of these.
- 9.27** List and describe briefly the microstructural changes that occur during overageing.

- 9.28 At any instant of time, the grain growth, i.e., the rate of increase of the grain radius with time, dr/dt , is proportional to the grain boundary energy stored per unit volume of the material. Show that the radius of spherical grains increases as the square root of time.

MULTIPLE CHOICE QUESTIONS

- For a spherical particle of radius r , the volume-to-surface area ratio is
A. $3/r$ B. $r/3$ C. $3r$ D. $\pi r/3$
- If the interfacial energy increases by 10%, the homogeneous nucleation barrier for a spherical particle increases by
A. 10% B. 21% C. 33% D. 100%
- If the nucleation barrier at 10°C of supercooling is 10^{-17} J, at 20°C of supercooling, it is
A. 2×10^{-17} J B. 4×10^{-17} J C. 2.5×10^{-18} J D. 5×10^{-18} J
- If the product phase completely wets a nucleating agent, the nucleation barrier as a fraction of homogeneous barrier is
A. 1 B. $\frac{1}{2}$ C. $\frac{1}{4}$ D. 0
- If the product phase does not wet at all the parent phase, the contact angle between the two phases is
A. 0° B. 45° C. 90° D. 180°
- When the contact angle is 60°, the heterogeneous nucleation barrier expressed as a fraction of the homogeneous barrier is
A. $\frac{1}{2}$ B. $\frac{1}{4}$ C. $\frac{1}{8}$ D. none of these
- As compared to the nucleation-rate maximum, the growth-rate maximum is at
A. a higher temperature
B. a lower temperature
C. the same temperature
D. the temperature of maximum transformation rate
- Fine grain sizes are obtained by
A. slow cooling B. increasing nucleation rate
C. decreasing growth rate D. fast cooling
- During pearlitic transformations
A. new phases form B. crystal structures change
C. compositions of phases change D. there is no diffusion

10. The hardness of martensite in a steel is a function of
A. C content B. cooling rate C. Ni content D. nose location
11. Martensitic transformations
A. are diffusion-controlled
B. are shear processes
C. yield two products of different compositions
D. yield a hard product in steels
12. The c/a ratio of martensite depends on the concentration of
A. Ni B. Mn C. C D. N
13. Bainite has
A. the same morphology as austenite
B. a non-lamellar morphology of ferrite and cementite
C. the coarsest morphology among all the products from austenite
D. none of these
14. During overageing, hardness
A. decreases B. increases
C. is constant D. increases abruptly
15. Overageing refers to
A. ageing above room temperature
B. ultrafine precipitate size
C. long ageing times
D. coarsening of precipitate particles
16. The maximum temperature up to which tungsten (m.p. = 3410°C) can be cold worked is approximately
A. 0°C B. 27°C C. 1200°C D. 1940°C
17. Lead melts at 327°C. It is hot rolled at
A. -273°C B. -200°C
C. room temperature D. none of these
18. The free energy decrease during recrystallization comes mainly from
A. excess point defects
B. excess dislocations
C. grain boundaries
D. lower energy of the new crystal structure
19. The recrystallization rate increases with
A. increasing amount of cold work
B. higher working temperature
C. higher annealing temperature
D. decreasing initial grain size
20. Grain growth occurs in the temperature range
A. $0.2-0.3 T_m$ B. $< 0.4 T_m$ C. $0.4-1.0 T_m$ D. $> T_m$

Answers

- | | | | | |
|----------|-------|------------|-------------|-------|
| 1. B | 2. C | 3. C | 4. D | 5. D |
| 6. D | 7. A | 8. B, C, D | 9. A, B, C | 10. A |
| 11. B, D | 12. C | 13. B | 14. A | 15. D |
| 16. C | 17. C | 18. B | 19. A, C, D | 20. C |

Sources for Experimental Data

ASM International, *Metals Handbook*, 10th ed., Vol. 1—*Properties and Selection: Iron, Steels and High Performance Alloys*, Materials Park, Ohio (1990).

W.D. Kingery, H.K. Bowen and D.R. Uhlmann, *Introduction to Ceramics*, Wiley, New York (1976), Chaps. 8–10.

Suggestions for Further Reading

W.D. Kingery, H.K. Bowen and D.R. Uhlmann, *Introduction to Ceramics*, Wiley, New York (1976), Chaps. 8–10.

V. Raghavan, *Solid State Phase Transformations*, Prentice-Hall of India, New Delhi (1987).

CHAPTER

10 Elastic, Anelastic and Viscoelastic Behaviour

Materials deform in response to an externally applied stress. This deformation can be permanent or temporary. Permanent deformation stays after the removal of the applied stress, while temporary deformation disappears on removal of the stress. Both temporary and permanent deformations can be functions of time or virtually independent of time. Permanent plastic deformation of inorganic crystalline materials is the subject matter of the next chapter. The time dependent part of the permanent deformation in crystalline materials called creep is also discussed there.

In this chapter, we first consider deformation which is temporary or recoverable. Fully recoverable deformation, which is virtually time independent, is called *elastic deformation*. Within the range of elastic behaviour, a distinction should be made between ordinary elasticity and rubber-like elasticity. Fully recoverable but time dependent deformation is called *anelastic deformation*. When both recoverable and permanent deformation occur together and are time dependent, we have *viscoelastic deformation*. This phenomenon is discussed at the end of this chapter.

Units

Quantity	SI units		Other units
	Unit	Symbol	
Interatomic force F	newton	N	kgf, lb, dyne
Applied force F_{app}			
Uniaxial stress σ			
Shear stress τ	meganeuton per square metre or megapascal	MN m^{-2} or MPa	kgf/mm^2 , psi, dyne/cm^2
Uniaxial strain ϵ			
Shear strain γ	—	—	—

Units (cont.)

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Young's modulus Y Shear modulus μ	giganewton per square metre or gigapascal	GN m^{-2}	kgf/mm^2 ,
		or	ksi, psi,
		GPa	dyne/cm^2
Entropy S	joule per mole per kelvin	$\text{J mol}^{-1} \text{K}^{-1}$	$\text{cal/gm}^\circ\text{C}$
Viscosity η	pascal second	Pa s	poise
Hysteresis loss	joule per cubic metre	J m^{-3}	—

* *Note:* Values of elastic moduli are given conveniently in units of GN m^{-2} ($\equiv 10^9 \text{ N m}^{-2}$). In this unit, a very low elastic modulus such as that of soft rubber is 0.001 GN m^{-2} ($\equiv 1 \text{ MN m}^{-2}$) and a very high modulus such as that of diamond is 1140 GN m^{-2} . GN m^{-2} can also be written as GPa. The Young's moduli of typical materials are listed in Appendices I and II.

ELASTIC BEHAVIOUR

Brittle materials, such as concrete, cast iron and silicate glasses, under tensile stress show elastic deformation right up to the point of fracture. Ductile materials such as copper and aluminium are elastic up to a certain stress called the elastic limit. Thereafter, they plastically deform. In both these groups of materials, within the elastic region, the strain is proportional to the stress applied, as given by Hooke's law. This behaviour can be called ordinarily elastic (or truly elastic) behaviour. The rubber-like behaviour, where the elastic strain is very large and is not a linear function of stress, is discussed in Sec. 10.3.

10.1 Atomic Model of Elastic Behaviour

The potential energy of a pair of atoms and the interatomic forces between them are shown as a function of their distance of separation r in Fig. 4.4. A general expression for the potential energy W is

$$W = -\frac{A}{r^n} + \frac{B}{r^m} \quad (10.1)$$

where A , B , n and m are constants and $m > n$. The negative term on the right is the attractive energy term. The positive energy term arises due to the mutual repulsion between the atoms at close distances of approach. The net interatomic force F is the negative of the derivative of W with respect to r :

$$F = -\frac{dW}{dr} \quad (10.2)$$

The potential energy is a minimum and the net interatomic force is zero at the equilibrium distance of separation r_0 , see Fig. 4.4.

In order to decrease or increase the interatomic distance of separation starting from the equilibrium value r_0 to r , an external force F_{app} equal in magnitude but opposite in sign to the interatomic force F at r must be applied.

$$F_{\text{app}} = -F = \frac{dW}{dr} \quad (10.3)$$

Referring to Fig. 10.1, to decrease the interatomic distance from r_0 to r' , a compressive force equal in magnitude to F' must be applied. To increase the

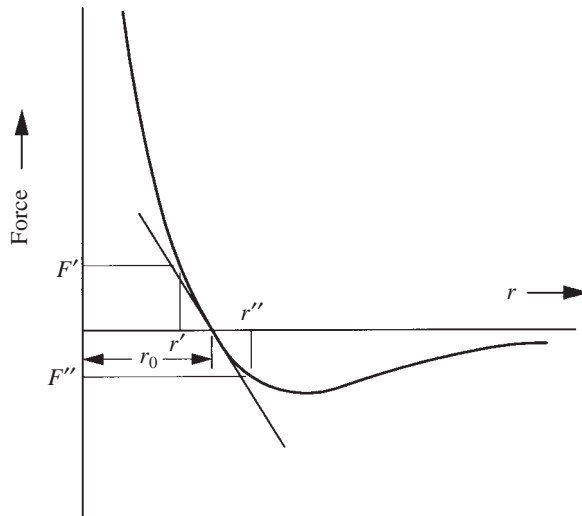


Fig. 10.1 The interatomic force-distance curve. Changing the bond length from r_0 to r' or r'' requires an external force of magnitude F' or F'' .

interatomic distance from r_0 to r'' , a tensile force equal in magnitude to F'' must be applied. A tangent to the force-distance curve drawn at r_0 practically coincides with the curve over a small range of displacement on either side of r_0 , see Fig. 10.1. Strains in the elastic region for both brittle and ductile materials lie in the range 0.001 to 0.005 and are, therefore, within this small range of displacement. Then, the negative of the slope of the force-distance curve at r_0 is proportional to the Young's modulus Y of a material. Also, the curvature of the potential energy curve at r_0 is proportional to the elastic modulus:

$$Y \propto -\frac{dF}{dr} = \frac{d^2W}{dr^2} \quad (10.4)$$

Example 10.1 Estimate the Young's modulus of a material, which has bonding characteristics as given by Eq. (10.1). Take that $n = 1$, $m = 9$, and $A = 7.68 \times 10^{-29}$ J m. The equilibrium distance between bonding atoms is 2.5 \AA .

Solution By setting the first derivative of the W - r function to zero, i.e.

$$\frac{dW}{dr} = \frac{A}{r^2} - \frac{9B}{r^{10}} = 0$$

we obtain

$$\begin{aligned} B &= A \times r^8/9 = 7.68 \times 10^{-29} \times 2.5^8 \times 10^{-80}/9 \\ &= 1.30 \times 10^{-106} \text{ J m}^9 \end{aligned}$$

Differentiating again, we have

$$\begin{aligned} \left. \frac{d^2W}{dr^2} \right|_{r=r_0} &= -\frac{2A}{r_0^3} + \frac{90B}{r_0^{11}} \\ &= -\frac{2 \times 7.68 \times 10^{-29}}{2.5^3 \times 10^{-30}} + \frac{90 \times 1.30 \times 10^{-106}}{2.5^{11} \times 10^{-110}} \\ &= 39.3 \text{ J m}^{-2} \end{aligned}$$

The curvature obtained above is also equal to the slope of the force-distance curve. To obtain the Young's modulus, we can take that the force is acting over an area of r_0^2 . So,

$$\begin{aligned} Y &= \frac{\text{stress}}{\text{strain}} = \frac{dF/r_0^2}{dr/r_0} \\ &= \frac{1}{r_0} \left(\frac{d^2W}{dr^2} \right)_{r=r_0} \\ &= 157 \text{ GN m}^{-2} \end{aligned}$$

Materials with strong bonds have a deep potential energy well with a sharp curvature. Hence, strong bonding results in large values for the elastic modulus. By the same token, the shallow potential well of the weakly bonded materials results in small values for the modulus. As we move to the right along a row of the periodic table, the covalent character of the bond and its strength increase. The Young's moduli of Li, Be, B and C (dia) in a row of the periodic table increase in that order.

Element	Li	Be	B	C (dia)
Atomic number Z	3	4	5	6
Young's modulus Y , GN m ⁻²	11.5	289	440	1140

Elements beyond carbon do not form solids with a three-dimensional network of covalent bonds. These would have moduli according to the strength of the secondary bonds in the solid, which are primarily stretched by an external stress. For the same reason, graphite, which consists of two-dimensional sheets held together by van der Waals bonds, has an average Young's modulus of only 8 GN m⁻², as compared to 1140 GN m⁻² for diamond, which has a three-dimensional network of primary bonds.

The covalent nature of the bonding decreases as we go down any given column of the periodic table. The elastic modulus shows a corresponding decrease. This trend is illustrated below for elements of the fourth column:

Element	C (dia)	Si	Ge	Sn	Pb
Atomic number Z	6	14	32	50	82
Young's modulus Y , GN m ⁻²	1140	103	99	52	16

In a crystal, the interatomic distance varies with crystal direction, with a corresponding variation in the bond strength. This gives rise to *elastic anisotropy*, that is, the elastic properties become a function of the crystal direction. This anisotropy is particularly evident in materials which have two kinds of bonds. For example, the Young's modulus of graphite in the a direction parallel to the sheets is 950 GN m⁻², which is much larger than that averaged over all directions, which is only 8 GN m⁻².

The transition metals have elastic moduli much higher than those of the alkali metals, as a result of the partial covalent character of their bonds. *Metals of the first transition series have Young's moduli in the range 200 GN m⁻²*. Metals of the second and the third transition series have higher moduli, reaching up to 600 GN m⁻². As is the case with some other physical properties, the modulus reaches a peak value for the electronic configuration of d^5 , see Fig. 4.9.

The application of a tensile stress causes an elongation ϵ_l along the tensile axis and a transverse contraction ϵ_t . The ratio of these two strains defines the *Poisson ratio* ν :

$$\nu = -\frac{\epsilon_t}{\epsilon_l} \tag{10.5}$$

The Poisson ratio for metals is around 0.3, for polymers and rubber it is between 0.4 and 0.5; for ionic solids, it is around 0.2.

The shear modulus of a material is defined as the ratio of the shear stress applied to the shear strain. The shear modulus μ is related to the Young's modulus Y through the Poisson ratio:

$$\mu = \frac{Y}{2(1 + \nu)} \tag{10.6}$$

Similarly, the bulk modulus K , which is the ratio of the hydrostatic stress to the relative volume change, is related to Y :

$$K = \frac{Y}{3(1 - 2\nu)} \tag{10.7}$$

The equilibrium distance of separation r_0 between atoms shown in Figs. 4.4 and 10.1 is applicable at 0 K, where there is no thermal energy. At higher temperatures, under the influence of thermal energy, the atoms vibrate about their mean positions, the amplitude of the vibrations increasing with increasing temperature. With more thermal energy, we can visualize the bonds to be somewhat loosened up. This reflects in a decrease in the elastic modulus with increasing temperature, see Fig. 10.2. In a majority of cases, on heating from 0 K to the melting point, the decrease in elastic modulus is in the range 10–20%.

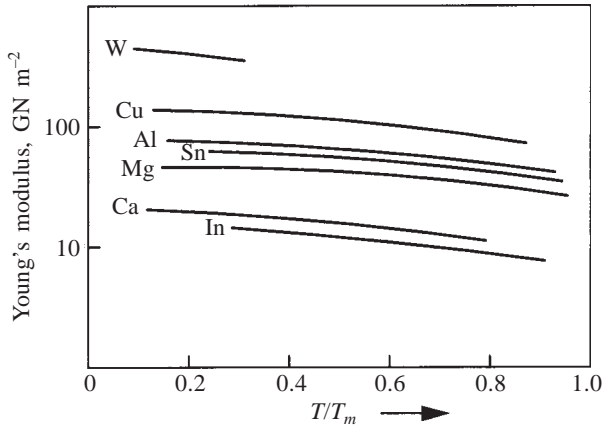


Fig. 10.2 Young's modulus for a few materials plotted against T/T_m .

10.2 The Modulus as a Parameter in Design

The *stiffness* of a material is its ability to resist elastic deformation or deflection on loading. The stiffness is dependent on the shape of the structural component. For identical shapes, it is proportional to the elastic modulus. Therefore, the elastic modulus is an important material parameter in mechanical design. Materials with high stiffness and hence high modulus are often required.

Covalently bonded elements such as diamond have a very high modulus (1140 GN m^{-2}). However, they are not suitable for use in engineering practice, due to high cost, nonavailability and brittleness. Brittle materials cannot withstand accidental overloading during service and may fail in a catastrophic manner. Hence, they are not suitable as structural members, even though they may have a high modulus.

Ductile elements such as metals withstand accidental overloading without catastrophic failure and as such are suitable for structural components. Among the metals, the elements of the first transition series offer a good compromise of adequate ductility and a moderately high modulus, in the range 200 GN m^{-2} . The metals of the second and the third transition series have an even higher modulus but have the disadvantage of high density. By suitable alloying, the Young's modulus of metals can be increased. However, the modulus being a

structure-insensitive property, it can be increased only in proportion to the concentration of the higher modulus additive. For producing a high modulus Fe-based material, reinforcement with TiB_2 is a promising route. With 50 vol% of TiB_2 particles in the Fe matrix, there is an increase of more than 50% in the modulus. The TiB_2 particles are in stable equilibrium with Fe. As the particles are approximately spherical, the modulus is not dependent on direction as in fibre-reinforced materials.

The Young's moduli of some *ionic solids* are given below.

Material	NaCl	MgO	Al_2O_3	TiC	Silica glass
Young's modulus Y , GN m^{-2}	37	310	402	308	70

Even though the modulus values of some of them are quite high, they also suffer from the lack of ductility like covalent solids.

In spite of their plasticity, *polymers* are not suitable for applications where high stiffness is required. They have a low modulus, as the chains are bonded together by secondary bonds. The value of the modulus is dependent on the nature of the secondary bonding (van der Waals or hydrogen bonding), the presence of bulky side groups, branching in the chains and cross-linking. For example, unbranched polyethylene has a Young's modulus of 0.2 GN m^{-2} , whereas polystyrene with a large phenyl side group in the monomer has a modulus of 3 GN m^{-2} . (Refer to Table 5.7 for comparison of the monomer structure.) Three-dimensionally bonded network polymers such as phenol formaldehyde and fully cross-linked rubber (ebonite) have a modulus in the range $3\text{--}5 \text{ GN m}^{-2}$. It is evident that polymers as a whole have much lower moduli, as compared to other primarily bonded materials. This places a severe restriction on the use of polymers as structural components.

In *composite materials*, an attempt is made to increase the stiffness, without the disadvantages of brittleness. Boron has a low density and is suitable for light weight applications and for air borne structures. Its elastic modulus is one of the highest for elements ($Y = 440 \text{ GN m}^{-2}$), but it is brittle. It can be used as a reinforcing fibre for a ductile matrix such as aluminium. In the Al-B composite, the elastic modulus is increased due to the presence of the boron fibres. At the same time, the disadvantages of the brittleness of boron are countered, by the cushioning effect of the ductile matrix. The ductile matrix stops a propagating crack if a fibre embedded in it breaks accidentally. If the entire material were to consist of boron only, a propagating crack would culminate in the fracture of the entire cross-section.

The Young's modulus Y_c of a composite in a direction parallel to the fibres can be expressed as a linear function of the moduli of the fibre and the matrix, Y_f and Y_m :

$$Y_c = V_f Y_f + V_m Y_m \quad (10.8)$$

where V_f and V_m are the volume fractions of the fibre and the matrix. Thus, a 40 vol.% of boron in an aluminium matrix can raise the Young's modulus from 71 GN m^{-2} for pure aluminium to 219 GN m^{-2} for the composite. This composite would then be as stiff as steel but less than one-third its density! The

volume fraction of the fibre in a composite material cannot be increased indefinitely, as at some stage the problems of aligning the fibres and keeping them separated from one another become serious.

Example 10.2 A tensile load of 100 N is applied to an aluminium-boron composite of 1 mm² cross-sectional area. The volume of the parallel fibres is 30%. What is the stress in the fibres, when the load axis is

- (i) parallel to the fibres, and
- (ii) perpendicular to the fibres?

Solution (i) When the load is applied parallel to the fibres, the elastic strain ε is the same in both the fibres and the matrix. If σ_f and σ_m are the stresses in the fibres and the matrix,

$$Y_f = \frac{\sigma_f}{\varepsilon_f}, \quad Y_m = \frac{\sigma_m}{\varepsilon_m}$$

Since $\varepsilon_f = \varepsilon_m$, we have

$$\frac{Y_f}{Y_m} = \frac{\sigma_f}{\sigma_m} = \frac{440}{71} = 6.20$$

Also, $0.3\sigma_f + 0.7\sigma_m = 100$ N, or

$$\sigma_m = \frac{100 - 0.3\sigma_f}{0.7} = 143 - 0.43\sigma_f$$

Then,

$$\frac{\sigma_f}{143 - 0.43\sigma_f} = 6.20$$

$$\sigma_f = 242 \text{ N mm}^{-2} (= 242 \text{ MN m}^{-2})$$

- (ii) When the load is applied perpendicular to the fibres, the stress in the fibre and the matrix is the same, i.e. 100 MN m⁻².
-

Figure 10.3 shows an aluminium-silica composite, where the aluminium-coated silica glass fibres have been partially pressed together to form the composite. One of the most popular composites comprises *silica glass fibres embedded in an epoxy resin matrix*. A typical thermosetting epoxy resin has a Young's modulus of 3 GN m⁻², while a silicate glass has a modulus of about 70 GN m⁻². So, the reinforcing effect is quite marked. The density of the composite being quite low, it can compete in terms of stiffness and strength for the same weight with some of the high strength steels. However, the cost of the fibre glass composite is on the high side.

Aligning molecules in the direction of high modulus is another method of increasing the stiffness in practical applications. The long chains in a polymer can be aligned by plastic working. When a stress is applied parallel to the chain length, the strong covalent bonds along the backbone of the chain tend to be stretched

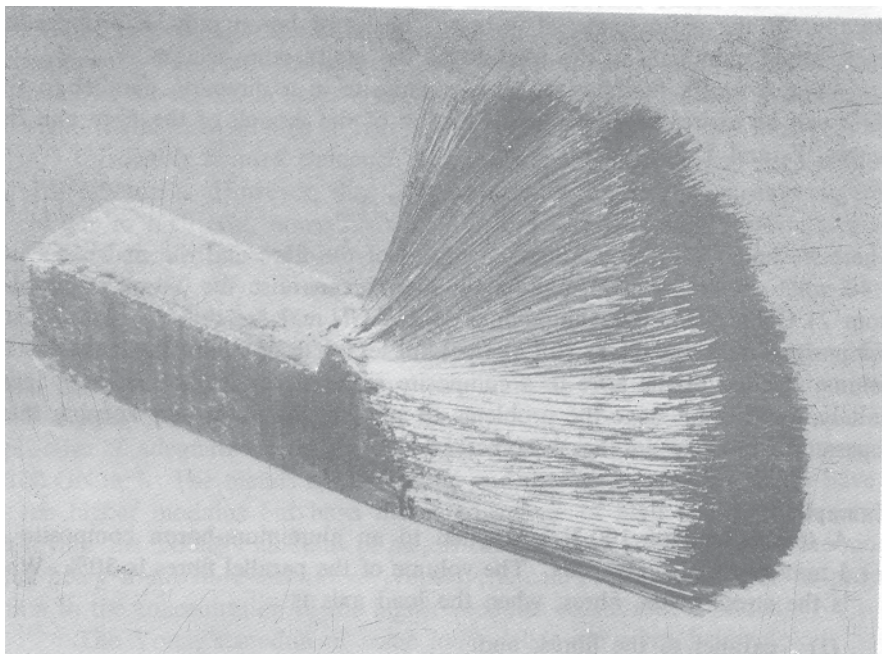


Fig. 10.3 Aluminium coated fibres of silica are pressed together to produce a strong and light composite material. (Courtesy: Rolls Royce Limited, Derby, UK.)

rather than the van der Waals bonds between chains. This increases the modulus of the aligned polymer. Polymeric fibres can have a modulus at least one order of magnitude higher than that of a non-aligned polymeric structure. By extending this argument, one would have expected that single crystals of a polymer should have a very high stiffness, when stressed in a direction parallel to the chains, as there is perfect alignment in a crystalline arrangement. Unfortunately, this has not been realized in practice, as the single crystals tend to have a folded chain structure. The fold tends to be a weak point in the parallel arrangement.

Some applications require a near-zero variation in elastic modulus with changes in ambient temperature. Alloys of iron with 36% Ni and 5% Cr have this property and are called *elinvars*. These alloys are used in tuning forks and radio synchronization, where an invariant modulus is required. Here, the usual decrease in modulus with increasing temperature is compensated by a slight decrease in the interatomic distance due to a magnetic effect. As may be expected, the magnetic effect also affects the thermal expansion. The related alloy of iron with 36% nickel called *invar* has zero coefficient of thermal expansion around room temperature.

10.3 Rubber-like Elasticity

Materials which undergo recoverable deformation of a few hundred per cent are called *elastomers* and exhibit rubber-like elasticity. The stress is not proportional to strain in these materials, in contrast to ordinary elastic materials. Elastomers

are long chain molecules with some *cross-linking* between the chains. This cross-linking is important, because this feature is what keeps the molecules from slipping past one another permanently during stretching. After cross-linking, the translational motion of chains gets restricted to segmental mobility between cross-linking points. When a stress is applied to an elastomer, equilibrium in the molecular configuration is established fairly quickly so that we can ignore the time dependent aspects of stretching as a first approximation.

In the unstretched state, the chain molecules are randomly coiled. A large number of configurations of equal potential energy are then possible. This large number of distinguishable arrangements means an appreciable *configurational entropy* and a low free energy. On application of an external stress, the coiled molecules respond by stretching out. The stretching reduces the number of possible configurations and hence lowers the configurational entropy. In the limit, when the molecules are all fully stretched out, the possible configuration is only one and the configurational entropy is zero, recall Eq. (2.5). When a stretched rubber is heated, the increase in thermal energy tends to coil back the uncoiled molecules against the stretching force.

The coiling and uncoiling of a long chain molecule in an elastomer is schematically illustrated in Fig. 10.4. Note that full stretching out does not mean

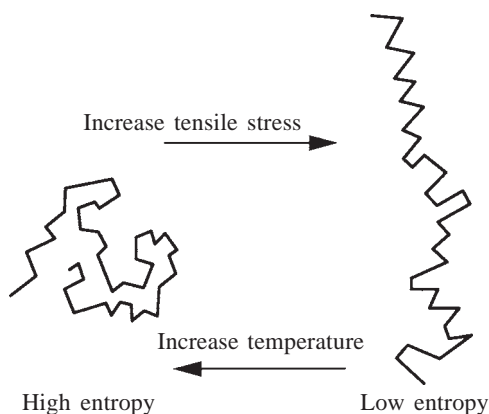


Fig. 10.4 The coiling and uncoiling of an elastomer chain molecule as a function of tensile stress and temperature.

a change in the C–C bond angle of 109.5° along the backbone of the chain. If this were to happen, the enthalpy (bond energy) of the elastomer would also change, in addition to the configurational entropy. Experimental results indicate that the change in the enthalpy on stretching a rubber is zero. The stretching process merely uncoils the coiled molecules, but does not change the bond lengths or bond angles. This behaviour is in contrast to what happens in an ordinary elastic material, where bond lengths are clearly changed, see Fig. 10.1. Using this experimental result, with the first and second laws of thermodynamics, it can be shown that the stretching force F at temperature T is related to the entropy S and length L of the material as follows:

$$F = -T \left(\frac{\partial S}{\partial L} \right)_T \quad (10.9)$$

The stretching is assumed to be a slow isothermal process, with equilibrium maintained at all times. From the Boltzmann's equation for configurational entropy (Eq. (2.5)), we can write for the entropy change in going from length L_0 to length L as

$$S - S_0 = k \ln \left(\frac{w}{w_0} \right) \quad (10.10)$$

where S_0 and w_0 are the entropy and the number of possible configurations before stretching and S and w refer to the same quantities after stretching. Using permutations, the entropy change can be expressed in terms of length as follows:

$$S - S_0 = -\frac{1}{2} N_0 k \left[\left(\frac{L}{L_0} \right)^2 + 2 \left(\frac{L_0}{L} \right) - 3 \right] \quad (10.11)$$

where N_0 is the number of chain segments between cross-linking points. Combining Eqs. (10.9) and (10.11), we get

$$F = \frac{N_0 k T}{L_0} \left[\left(\frac{L}{L_0} \right) - \left(\frac{L_0}{L} \right) \right] \quad (10.12)$$

Note that, at constant temperature, the quantity outside the square brackets on the right-hand side is constant. This relationship can be experimentally verified, by merely noting the force required to obtain a certain change in length.

Equation (10.12) is the equation of state of the rubbery material. This can be written as

$$\phi(F, L, T) = 0 \quad (10.13)$$

If this function ϕ is continuous and single valued, there is a theorem of differential calculus which states that:

$$\left(\frac{\partial F}{\partial L} \right)_T \left(\frac{\partial L}{\partial T} \right)_F \left(\frac{\partial T}{\partial F} \right)_L = -1 \quad (10.14)$$

Equation (10.14) can also be verified experimentally, by obtaining the incremental slopes corresponding to each of the three partial differentials.

The stress-strain curve for an elastomer is shown in Fig. 10.5. The dotted curve is calculated from Eq. (10.12). The experimental curve follows the dotted curve up to a large amount of strain. The steep increase in the slope of the experimental curve in the later stages can be attributed to bond stretching in the straightened-out molecules. Bond stretching is not accounted for in Eq. (10.12).

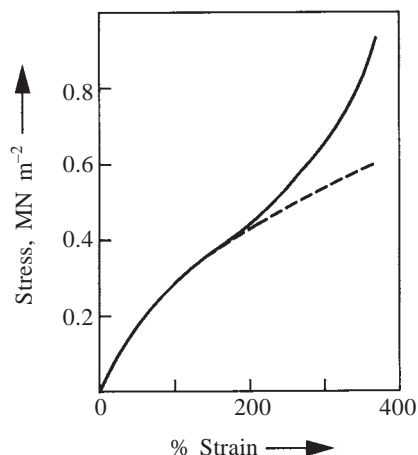


Fig. 10.5 The stress-strain curve for an elastomer.

ANELASTIC BEHAVIOUR

Recoverable deformation that occurs as a function of time is called anelastic deformation. The elastic deformation of the material continues even after the application of the load, due to some relaxation process within the material. On removal of the load, some part of the elastic deformation is recovered only as a function of time, with the reversal of the relaxation process.

10.4 Relaxation Processes

Several relaxation processes take place within a material in response to an externally applied stress. If the time scale of a *relaxation process* is *too fast or too slow* compared to the time interval over which the stress is applied, the stress-strain relationship is essentially independent of time. If, however, the time scale of the process is comparable to the time interval of stress application, the stress-strain relationship is *dependent on time* and results in *anelastic behaviour*.

The fraction X of a relaxation process that has been completed up to time t can be described by an exponential relation:

$$X = 1 - \exp \left(-\frac{t}{t_r} \right) \quad (10.15)$$

where t_r is called the *relaxation time*. When $t = t_r$, $X = 0.63$. The relaxation time t_r is then the time taken for 63% of the relaxation to take place and is a useful parameter to specify the time scale of a relaxation process.

Consider the following relaxation process in a dilute solution of carbon in BCC (α) iron. The interstitial carbon atoms stay at the middle of the cube edges (which are $\langle 100 \rangle$ directions) in the BCC unit cell of iron. The largest sphere that will fit in this interstitial position has a radius of about 0.2 \AA . This is very much smaller than the radius of the carbon atom, 0.71 \AA . A carbon atom in this interstitial position tends to push apart the two iron atoms on either side of it, resulting in a large distortional energy. Consequently, only a very small percentage of carbon (0.008%) goes into solution in BCC iron.

On applying a tensile stress, one of the three $\langle 100 \rangle$ directions approximately parallel to the tensile axis elongates, while the other two axes perpendicular to it contract due to the Poisson effect. The carbon atoms along the two contracting axes jump to occupy vacant positions along the elongated axis, as there is more space available there. This reduces the total distortional energy around the interstitial atoms. This jumping results in an additional stretching in the direction of the applied stress. The jumping of the carbon atoms is a time dependent diffusion process with a relaxation time $t_r \sim 100 \text{ s}$ at room temperature.

Let the BCC iron with a small quantity of carbon in solution be subjected to a number of alternating cycles of loading and unloading within the elastic region. If the time taken for each cycle is very small compared to the relaxation time given above, the carbon atoms will not have enough time to jump from the contracted axis to the elongated axis before the stress reversal takes place. Under such conditions, the carbon atoms do not jump at all. The stress-strain curve (Fig. 10.6a) simply corresponds to bond stretching, which is instantaneous and accounts for the ordinary elastic behaviour.

If the time taken for stress reversals during cyclic loading is very large compared to the relaxation time for carbon jumping, the carbon atoms will redistribute themselves readily as loading or unloading proceeds. At every stage during loading or unloading, the additional strain due to the preferential occupation by carbon atoms of sites along the tensile axis will add to the strain due to bond stretching. The slope of the stress-strain curve will be correspondingly smaller, Fig. 10.6b, as compared to the case in Fig. 10.6a, where no redistribution occurs.

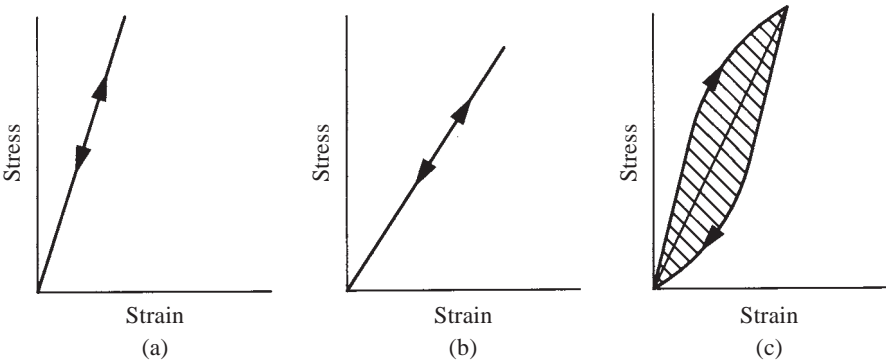


Fig. 10.6 The stress-strain curve under cyclic loading at different frequencies.

Finally, consider the situation where the time taken for each cycle of loading and unloading is about the same as the relaxation time. The carbon jumping will continue to occur, as loading is done. The strain due to carbon jumping will somewhat lag behind the strain due to bond stretching, which is instantaneous. Even after the maximum load has reached, the strain due to carbon jumping will continue to occur, resulting in further strain as a function of time. Due to this effect, the stress-strain curve during loading does not coincide with the curve during unloading, which is lower as illustrated in Fig. 10.6c. As the cyclic loading is continued, the strain continues to lag behind stress, so that a hysteresis loop is traced during each cycle. The elastic strain energy equal to the area under the loop is lost during each cycle.

Example 10.3 Estimate the diffusion coefficient D for carbon at room temperature, using the experimental value of the relaxation time $t_r \sim 100$ s.

Solution From Eq. (8.16), $D = v'\delta^2$, where δ is the jump distance and v' is the jump frequency of carbon atoms. As the relaxation time $t_r \sim 100$ s, the jump frequency is $1/100 = 0.01 \text{ s}^{-1}$. Taking the jump distance to be 2.5 \AA , the diffusion coefficient $D = 2.5^2 \times 10^{-20} \times 0.01 = 6.25 \times 10^{-22} \text{ m}^2 \text{ s}^{-1}$. The value calculated from the data in Table 8.2 is $2.8 \times 10^{-21} \text{ m}^2 \text{ s}^{-1}$, agreeing with the above result in order of magnitude.

Relaxation processes in materials have a very wide range of time spectrum from 10^{-13} second for atomic vibrations to more than 10^6 seconds for the viscous flow of grain boundaries. For a given process, the relaxation time is strongly dependent on temperature and decreases with increasing temperature in an exponential fashion. A relaxation process that does not give rise to hysteresis losses at a certain temperature may do so at some other temperature, as the relaxation time changes on heating or cooling, in relation to the frequency of the loading cycle.

It is important that, during anelastic deformation, the heat generated during each cycle of loading equal to the area of the hysteresis loop should be dissipated as soon as possible. The fraction of this energy that is dissipated is dependent on the nature of the material and is called the *damping capacity* of the material. Gray cast iron has a good damping capacity as compared to mild steel. This can be attributed to the presence of graphite flakes in the microstructure of gray cast iron, see Fig. 7.14. It is, therefore, used as a base for erection of machinery to damp out vibrations efficiently. Polymers including rubber have very good damping capacity compared to other materials and find uses such as mounting pads for delicate instruments.

A low damping capacity is desirable only in some special applications as in temple bells and piezoelectric devices. For example, the *quartz crystal* is used as ultrasonic flaw detector and for obtaining elastic constants. This is done by measuring the time interval for the echo from an ultrasonic wave traversing up

and down the crystal. Here, the time between two successive echos is measured and a low damping capacity aids in delineating the echo peaks, before appreciable decay occurs.

For a given relaxation process, the energy loss during a cycle is obviously a maximum at some frequency of loading. In automobile tyres, the frequency of loading increases with the speed of the vehicle. In the common type of tyres, the energy loss becomes a large value at high speeds such as 100 km per hr. This would increase the heating effect in the tyres, reducing their lifetime as well as increasing the risk of a tyre burst at high speeds. Recent improvements such as fibre-glass reinforced polyester tyres yield a much larger relaxation time than the unreinforced tyres for the same temperature, so that tyre life and the safety factor are increased.

VISCOELASTIC BEHAVIOUR

Viscoelastic behaviour is found in materials which respond to an applied stress by both recoverable and permanent deformations, which are time dependent. This behaviour is very common in noncrystalline organic polymers. Time-dependent permanent deformation is called viscous flow. It is analogous to the creep phenomenon in crystalline materials discussed in Chapter 11.

10.5 Spring-Dashpot Models

Macroscopic mechanical models are often used to describe the viscoelastic behaviour of a material. Combinations of the classical *linear laws of Hooke and Newton* are the bases for these models. The shear strain γ is proportional to the shear stress τ in Hooke's law:

$$\tau = \mu \gamma \quad (10.16)$$

where μ is the shear modulus. Newton's law for viscous flow states that the (viscous) shear strain rate ($d\gamma/dt$) is proportional to the stress applied:

$$\tau = \eta \left(\frac{d\gamma}{dt} \right) \quad (10.17)$$

where η is the viscosity of the material. A *spring* represents the perfectly elastic behaviour. A *dashpot* represents a pure viscous behaviour. The dashpot consists of a container with a piston moving in a viscous liquid.

The *Maxwell element* is a series combination of a spring and a dashpot, as depicted in Fig. 10.7a. The spring and the dashpot support the same stress. The strain in the element is the sum of the strain in the spring and the strain in the dashpot:

$$\gamma = \gamma_{\text{spring}} + \gamma_{\text{dashpot}} \quad (10.18)$$

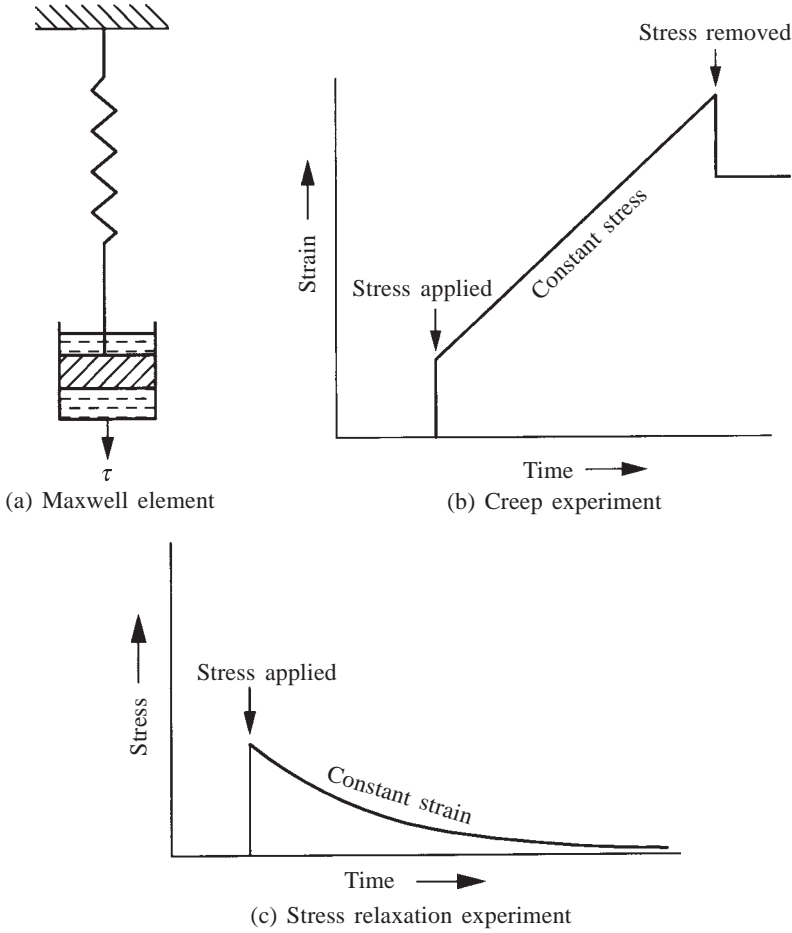


Fig. 10.7 The creep and the stress-relaxation experiments on a Maxwell element.

Differentiating Eq. (10.18) with respect to time, we have

$$\frac{d\gamma}{dt} = \left(\frac{d\gamma}{dt} \right)_{\text{spring}} + \left(\frac{d\gamma}{dt} \right)_{\text{dashpot}} \quad (10.19)$$

Noting that $(d\gamma/dt)_{\text{spring}} = (d\tau/dt)/\mu$ and $(d\gamma/dt)_{\text{dashpot}} = \tau/\eta$, we obtain, for the stress on the element, the equation

$$\tau = \eta \left(\frac{d\gamma}{dt} \right) - \frac{\eta}{\mu} \left(\frac{d\tau}{dt} \right) = \eta \left(\frac{d\gamma}{dt} \right) - t_r \left(\frac{d\tau}{dt} \right) \quad (10.20)$$

where the relaxation time $t_r = \eta/\mu$.

The Maxwell element can be subjected to two different kinds of experiments. The first is called the *creep test*. Here, a stress is applied

instantaneously; the spring responds, giving an instantaneous strain. Then, as a function of time, the dashpot responds, increasing progressively the viscous strain. After a certain time interval, the stress can be removed. The spring then responds by immediate recovery of the elastic strain. This is illustrated in Fig. 10.7b. In the other type of experiment called the *stress relaxation test*, the stress is applied instantaneously and the strain thereafter is kept constant. The stress on the element decreases as a function of time, as the elastic strain is replaced by plastic strain induced by the viscous flow of the dashpot (see Fig. 10.7c). The Maxwell element thus exhibits instantaneous elastic deformation and elastic recovery (like a solid) and viscous flow (or creep) as a function of time (like a liquid).

In the *Voigt–Kelvin element*, the spring and the dashpot are connected in parallel, as shown in Fig. 10.8a. Then the strain in both of them is the same at any time:

$$\gamma = \gamma_{\text{spring}} = \gamma_{\text{dashpot}} \quad (10.21)$$

The stress supported by the element is the sum of the stresses in the spring and the dashpot:

$$\tau = \tau_{\text{spring}} + \tau_{\text{dashpot}} \quad (10.22)$$

Combining Eq. (10.22) with the Hooke's and the Newton's laws, we obtain

$$\tau = \mu\gamma + \eta\left(\frac{d\gamma}{dt}\right) \quad (10.23)$$

On instantaneous application of the stress, the dashpot offers a high initial resistance, as it cannot respond instantaneously. With time, the strain in the element increases, but the spring offers an increasing amount of resistance so that the strain rate decreases progressively. Eventually, the system attains equilibrium, with only the spring supporting the entire stress. On removal of stress, the strain decays exponentially. This is illustrated in Fig. 10.8b. In the

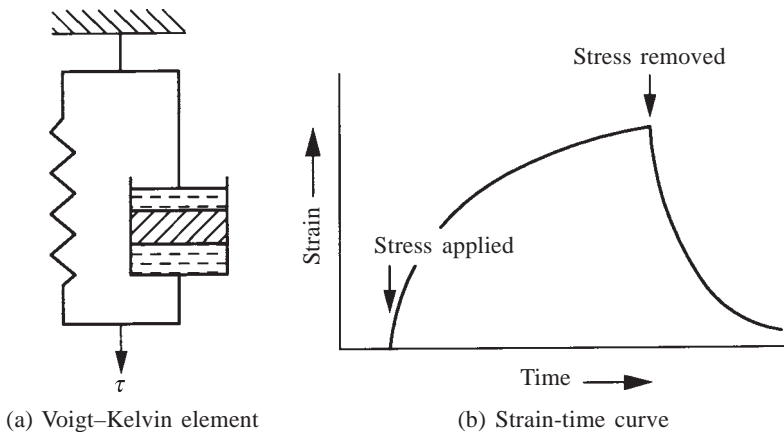


Fig. 10.8 The strain-time relationship for a Voigt–Kelvin element.

Voigt–Kelvin element, the strain always changes only as a function of time and not instantaneously. There is no permanent deformation at the end of the experiment.

Consider a linear combination of a Maxwell element and a Voigt–Kelvin element. This is called a *four parameter model*, as four parameters (two viscosities and two shear moduli) should be known to describe the exact behaviour of a material, that is modelled on this basis. The spring of the Maxwell element would represent the stretching of bonds in the material, which is instantaneous and fully elastic. The dashpot of the Maxwell element would represent pure viscous flow, such as slippage of chain molecules past one another permanently. The Voigt–Kelvin element would represent a time dependent but recoverable deformation, such as the coiling and the uncoiling of chain molecules. Note that we ignored the time dependence of this process, when we considered the rubber-like elasticity. The time dependent uncoiling of molecules on application of a stress is called *retarded elasticity*. The time dependent coiling on removal of the stress can be called the *entropy elasticity*, as it is due to the restoring force of the thermal agitation to revert the molecules to the coiled, high-entropy state. Both the coiling and uncoiling of molecules can be treated as relaxation processes that give rise to anelastic behaviour.

SUMMARY

1. The response of materials to an externally applied stress can be recoverable or permanent deformation or both. Each of these deformations can be time dependent or independent.
2. Ordinary elastic behaviour can be attributed to the instantaneous stretching of atomic bonds. The stress is proportional to strain here, as given in Hooke's law. Strongly bonded materials have high modulus, while weakly bonded materials have low modulus.
3. Several methods are available to increase the elastic modulus of a material, such as making composites and aligning chain and sheet molecules in directions of high bond strength.
4. Rubber-like elasticity arises from entropy changes during stretching, when the coiled long chain molecules uncoil.
5. The time dependent part of recoverable deformation is always associated with a relaxation phenomenon. The relaxation processes in materials are numerous, covering a wide range of time scale.
6. Energy losses due to anelastic processes should be kept a minimum and should be dissipated efficiently in most applications.
7. Long chain polymers exhibit viscoelastic behaviour. Using combinations of macroscopic models due to Maxwell and Voigt–Kelvin, the entire spectrum from purely elastic to fully viscous behaviour of polymers can be described.

PROBLEMS

- 10.1** The Young's moduli of alkali metals are given below in units of GN m^{-2} : Li (11.5), Na (9.0), K (3.5), Rb (2.7), Cs (1.8) and Fr (1.7). Compare this with the corresponding values of the melting points. Give a reason for this sequence.
- 10.2** Arrange the Young's moduli of the elements from Na to Si in order of increasing atomic number and compare it with the trend from Li to C (dia). Comment on your findings.
- 10.3** Plot the Young's modulus of the elements of the first, the second and the third transition series as a function of their atomic number, using the data from Appendix I. Comment on your findings.
- 10.4** The shear modulus of graphite averaged over all crystal directions is 8 GN m^{-2} . How can graphite be used to reinforce titanium, which has a shear modulus of 40 GN m^{-2} ?
- 10.5** A steel bar and an aluminium bar are each under a load of 5000 N. If the cross-sectional area of the steel bar is 100 mm^2 , what must be the area of aluminium for the same elastic deformation?
Answer: 290 mm^2 .
- 10.6** A tensile load of 5000 N is applied to a plastic-glass fibre composite of 100 mm^2 cross-sectional area, with 25 vol% of glass fibre. Calculate the cross-sectional area of the plastic alone (without reinforcement) to withstand the same load with the same elastic deformation.
 Given: $Y_{\text{plastic}} = 0.35 \text{ GN m}^{-2}$ and $Y_{\text{fibre}} = 70 \text{ GN m}^{-2}$
Answer: 5075 mm^2 .
- 10.7** A fibre-reinforced polystyrene contains 75 wt% of borosilicate glass fibre-oriented in a parallel fashion. Estimate the Young's modulus of the composite in the longitudinal direction of the fibres. Given:
- | <i>Property</i> | <i>Borosilicate</i> | <i>Polystyrene</i> |
|-------------------------------------|---------------------|--------------------|
| Young's modulus, GN m^{-2} | 65 | 2.6 |
| Specific gravity | 2.4 | 1.05 |
- Answer:* 38 GN m^{-2} .
- 10.8** Explain why an elastomer under tensile load contracts on heating. Compare this with the behaviour of ordinary materials.
- 10.9** Does the thermodynamic analysis of rubber elasticity apply to rubber in compression? Explain.
- 10.10** The Young's modulus of an elastomer increases from 0.001 GN m^{-2} to 1 GN m^{-2} on cross-linking. Explain why this is so.
- 10.11** Elevated temperatures are required to observe the relaxation due to viscous sliding of grain boundaries. Why?

- 10.12** Derive the result given in Eq. (10.20).
- 10.13** Using Eq. (10.20), show that the stress τ as a function of time t at constant strain in the stress-relaxation experiment decays exponentially as $\tau = \tau_0 \exp(-t/t_r)$, where τ_0 is the initial stress.
- 10.14** Sketch the time-strain curve for anelastic behaviour and compare it with the time-strain curve in the Voigt–Kelvin element. What do you deduce from this?
- 10.15** Explain how the four constants corresponding to the linear combination of the Maxwell and the Voigt–Kelvin elements can be determined from a strain-time experiment at constant stress.
- 10.16** Calculate the Young's modulus of a composite containing 60 vol% of glass fibre ($Y = 70 \text{ GN m}^{-2}$) in a matrix of epoxy resin ($Y = 3 \text{ GN m}^{-2}$) under isostress conditions.
Answer: 7.0 GN m^{-2} .

MULTIPLE CHOICE QUESTIONS

1. High elastic modulus in materials arises from
 - A. High strength of bonds
 - B. Weak bonds
 - C. Sharp curvature at the minimum potential energy
 - D. Shallow potential well
2. As we go along a row of the periodic table, if the three-dimensional network of primary bonds persists, the elastic modulus

A. increases	B. remains constant
C. decreases	D. decreases sharply
3. As we go down a column of the periodic table, the elastic modulus

A. increases sharply	B. increases
C. remains constant	D. decreases
4. The change in Young's modulus of ordinarily-elastic materials between 0 K and melting point is

A. 10–20% increase	B. no change
C. 10–20% decrease	D. 80–90% decrease
5. Covalent and ionic solids are not suitable as structural components, because

A. they have weak bonds	B. they are ductile
C. they are brittle	D. they have high elastic moduli

6. During stretching of an ideal elastomer, its enthalpy
 - A. increases
 - B. remains constant
 - C. decreases
 - D. decreases slowly
7. On heating an elastomer under tensile load, it shrinks
 - A. to maximize the enthalpy
 - B. to maximize the entropy
 - C. to minimize the free energy
 - D. to avoid breaking
8. On heating a rubber under a tensile force, it
 - A. shrinks
 - B. expands
 - C. expands rapidly
 - D. shows no change
9. In BCC iron, residual carbon occupies the midpoints of
 - A. $\langle 100 \rangle$
 - B. $\langle 110 \rangle$
 - C. $\langle 111 \rangle$
 - D. none of these
10. During cyclic loading, hysteresis loss occurs due to anelastic deformation, when
 - A. time for stress reversal is large compared to relaxation time
 - B. time for stress reversal is small compared to relaxation time
 - C. time for stress reversal and the relaxation time are about the same
 - D. none of these
11. The Voigt–Kelvin element is a
 - A. series combination of a spring and a dashpot
 - B. parallel combination of a spring and a dashpot
 - C. a four-parameter model
 - D. none of these
12. A linear combination of a Maxwell and a Voigt–Kelvin element is a
 - A. one parameter model
 - B. two parameter model
 - C. three parameter model
 - D. four parameter model

Answers

- | | | | | |
|---------|---------|------|------|-------|
| 1. A, C | 2. A | 3. D | 4. C | 5. C |
| 6. B | 7. B, C | 8. A | 9. A | 10. C |
| 11. B | 12. D | | | |

Sources for Experimental Data

K.A. Gschneidner Jr., in *Solid State Physics* (1964), Vol. 16, p. 275 on *Physical Properties and Interrelationships of Metallic and Semimetallic Elements*.

R.C. Progelhof and J.L. Throne, *Polymer Engineering Principles*, Carl Hanser Verlag, Munich (1993).

Suggestions for Further Reading

N.G. McCrum, C.P. Buckley and C.B. Bucknall, *Principles of Polymer Engineering*, Oxford University Press, Oxford (1988).

D. Rosenthal, *Resistance and Deformation of Solid Media*, Pergamon Press, New York (1974), Chaps. 1 and 3.

CHAPTER 11 Plastic Deformation and Creep in Crystalline Materials

The permanent deformation of materials on the application of a load can be either plastic deformation or creep. In crystalline materials, at temperatures lower than $\sim 0.4T_m$, where T_m is the melting point in kelvin, the permanent deformation is called *plastic deformation*. In this temperature range, the amount of deformation that occurs after the application of load is small enough to be ignored. The rate at which the material is deformed may, however, play a role in determining the deformation characteristics. At temperatures above $\sim 0.4T_m$, permanent deformation continues as a function of time, following the application of the load. This behaviour is termed *creep* and is discussed at the end of this chapter.

Units

Quantity	SI units		Other units
	Unit	Symbol	
Shear stress τ	meganewton square metre	MN m ⁻²	psi, ksi, kgf/mm ² , tons/sq. in.
Tensile stress σ			
Yield strength σ_y			
Ultimate tensile strength (UTS)			
Strain ϵ	—	—	—
Strain rate ($d\epsilon/dt$)	per second	s ⁻¹	—
Shear modulus μ	giganewton per square metre	GN m ⁻²	psi, kgf/mm ² , dyne/cm ²
Dislocation velocity v_d	metre per second	m s ⁻¹	—
Dislocation density ρ	per square metre	m ⁻²	per sq. inch.
Grain diameter d	millimetre, micrometre	mm, μ m	inch
Hall–Petch constant k	meganewton per (metre) ^{3/2}	MN m ^{-3/2}	kgf/mm ^{3/2}

*Note: Yield strengths and tensile strengths are conveniently written in units of MN m⁻² ($\equiv 10^6$ N m⁻²). In this unit, a low strength such as that of a linear chain polymer is about 20 MN m⁻² (2 kgf/mm²), while that of a high strength alloy steel is about 2000 MN m⁻² (200 kgf/mm²). MN m⁻² can also be written as MPa or N mm⁻².

PLASTIC DEFORMATION

11.1 The Tensile Stress-Strain Curve

Plastic deformation can occur under tensile, compressive or torsional loads. A typical tensile stress-strain curve is shown in Fig. 11.1. The applied load is plotted against the elongation or extension of the test specimen in Fig. 11.1a. The applied load P divided by the initial cross-sectional area A_0 of the specimen gives the engineering stress:

$$\text{Engineering stress} = \frac{P}{A_0} \quad (11.1)$$

Engineering strain is given by the fractional increase in the gauge length l_0 :

$$\text{Engineering strain} = \frac{\Delta l}{l_0} \quad (11.2)$$

where Δl is the increase in gauge length.

The deformation of the specimen is elastic up to the yield point, beyond which it becomes plastic. The load at the yield point divided by the initial cross-sectional area of the test specimen is called the yield stress or the *yield strength* of a material. Mild steel exhibits a clearly defined yield point, but a number of other materials do not have a clear demarcation between the elastic and the plastic regions.

Beyond the yield point, the linear elastic region is followed by a nonlinear plastic region. In this region, the load required to cause further deformation increases with increasing strain. This phenomenon is called *work hardening*. The slope of the load-elongation curve decreases, as the elongation increases. It becomes zero at some maximum load. The engineering stress corresponding to the maximum load is called the *ultimate tensile strength* (UTS) of the material. Beyond this maximum, a *neck* forms somewhere in the middle of the specimen, where the cross-sectional area *locally decreases*. The applied load decreases up to the point of fracture, where the specimen breaks into two pieces across the reduced cross-section of the neck.

Figure 11.1a can be replotted in terms of true stress and true strain. True stress is the actual stress, which is the applied load divided by the minimum cross-sectional area A_i of the specimen at any instant:

$$\text{True stress} = \frac{P}{A_i} \quad (11.3)$$

When the increase in length dl is expressed as a fraction of the length l at any instant, we obtain the true strain:

$$\text{True strain} = \int_{l_0}^l \frac{dl}{l} \quad (11.4)$$

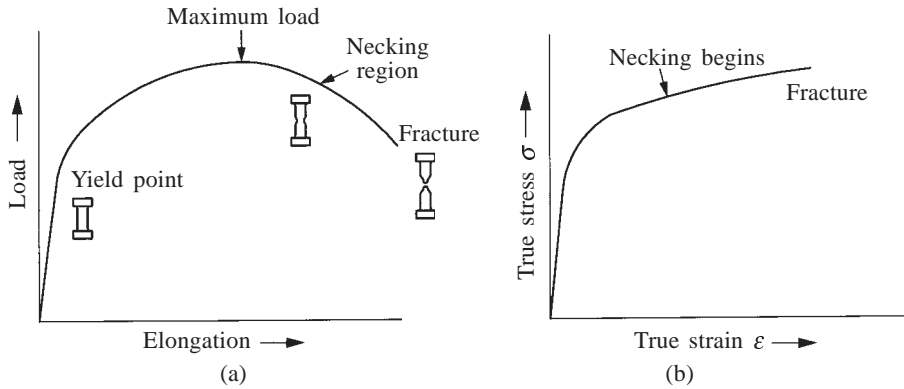


Fig. 11.1 The tensile load-elongation curve and the true stress-true strain curve for a ductile material.

When the deformation of the specimen becomes nonuniform after necking starts, the true strain becomes a function of the length over which it is measured. In order to avoid ambiguity, it is specified as the integral of $-dA/A$, where A is the cross-sectional area at the neck. The true stress-true strain curve is plotted in Fig. 11.1b. Unlike the load-elongation curve, there is no maximum in the true stress-true strain curve. The slope in the plastic region decreases with increasing strain, but does not become zero before fracture. This indicates that there is no work softening beyond the maximum in the load-elongation curve. The material work hardens continuously till fracture, although at a decreasing rate.

A power relationship is often used to express the true stress σ as a function of true strain ϵ :

$$\sigma = K\epsilon^n \quad (11.5)$$

where K is called the strength coefficient and n is the work hardening exponent. Materials which have a high work hardening exponent such as copper and brass ($n \sim 0.5$) can be given a large plastic strain more easily than those which have a smaller n , such as heat treated steel ($n \sim 0.15$).

Time does not enter as a parameter in Fig. 11.1. Ordinarily, the deformation characteristics can be taken to be independent of time. If, however, the rate of straining is varied over several orders of magnitude, the stress-strain curve may change as a function of strain rate ($d\epsilon/dt$). Another power relationship can be used to express σ at a given strain ϵ , in terms of the strain rate ($d\epsilon/dt$):

$$\sigma = A \left(\frac{d\epsilon}{dt} \right)^m \quad (11.6)$$

where A is a constant and m is the index of strain rate sensitivity. If $m = 0$, the stress is independent of the strain rate and the stress-strain curve would be the same for all strain rates. $m \sim 0.2$ for common metals. If $m = 0.4$ – 0.9 , the material may exhibit *superplastic behaviour*, that is, deform by several hundred per cent of strain without necking. The reason for this is that, as soon as necking

starts in some region, the strain rate increases locally, resulting in a rapid increase of the stress required to cause further deformation in that region. The deformation then shifts to another region of the material, where there is no necking. Here, the strain rate and hence the stress to cause deformation are smaller. Some stainless steels and aluminium alloys with a very fine grain size exhibit superplastic behaviour. The glass blower is able to pull his working material to very long rods without necking, because the exponent m for glass approaches one. If $m = 1$, the material behaves like a viscous liquid and exhibits Newtonian flow, see Eq. (10.17).

11.2 Plastic Deformation by Slip

X-ray diffraction studies show that the crystalline order in the solid is not lost during plastic deformation, even though more imperfections are introduced. The atom movements are such that the crystal structure remains the same before and after plastic deformation.

There are two basic modes of plastic deformation called slip and twinning. Slip is a shear deformation that moves atoms by many interatomic distances relative to their initial positions, as illustrated in Fig. 11.2b. Steps are created at the surface of the crystal during slip, but the orientation of all parts of the crystal remains the same before and after slip. *Twinning*, on the other hand, changes the orientation of the twinned parts, see Fig. 11.2c. Here, the movement of an atom

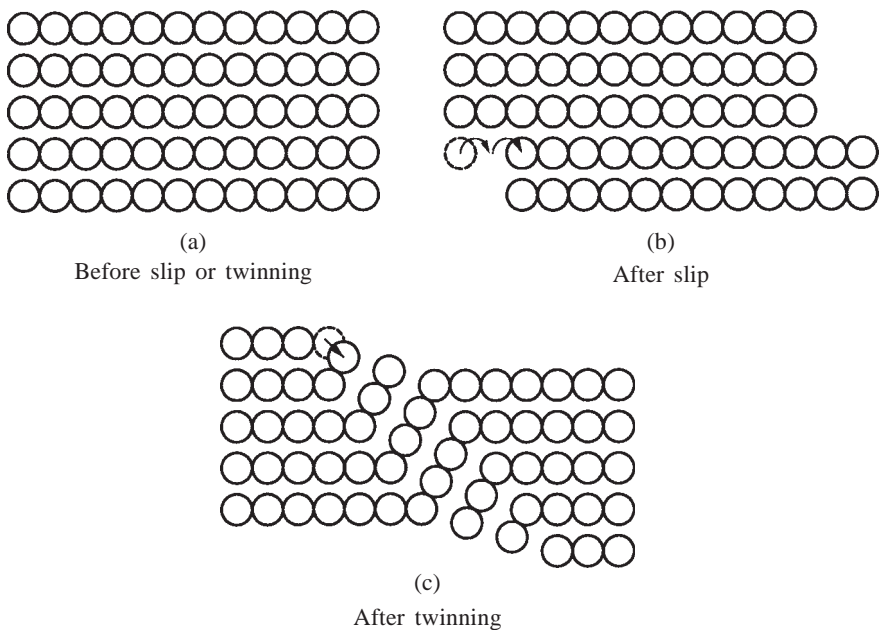


Fig. 11.2 The slip mode and the twinning mode of plastic deformation.

relative to its neighbours is only a fraction of an interatomic distance. The slip mode of deformation is the common mode in many crystals at ambient and elevated temperatures. At low temperatures, the mode of deformation changes over to twinning in a number of cases. We shall consider deformation by slip only.

Careful examination of the surface of a deformed crystal under the microscope shows groups of parallel lines, which correspond to steps on the surface. They are called *slip lines*. They indicate that the atomic planes within the crystal have sheared with respect to one another resulting in the surface steps. It is generally found that the slip planes are the closest packed planes in the crystal. Also, the directions along which slip occurs are the closest packed directions. It turns out that the planes of the greatest atomic density, having the highest number of atoms per unit area, are the most widely spaced planes. The directions of the greatest linear atomic density (close packed directions) have the smallest translation distance from one minimum energy position to the next. In ionic crystals, the slip planes and the slip directions are such that the ions of the same polarity do not become juxtaposed as nearest neighbours during shear, as this would mean a big increase in the potential energy of the crystal. The common slip planes and slip directions for some simple crystals are given in Table 11.1.

TABLE 11.1
Slip Planes and Directions in Some Crystals

Crystal		Slip planes	Slip directions
FCC		{111}	$\langle 110 \rangle$
BCC	more common	{110}	$\langle 111 \rangle$
	less common	{112}, {123}	
NaCl		{110}	$\langle 110 \rangle$
HCP	more common	basal plane	close packed directions
	less common	prismatic and pyramidal planes	

A slip plane and a slip direction that lies on it together constitute a slip system. For example, the combination (111) and $[\bar{1}10]$ forms a slip system, but not (111) and [110], as the [110] direction does not lie on the (111) plane. Counting the slip systems for the most densely packed slip planes only, there are 12 slip systems in FCC and BCC crystals, while there are only 3 in HCP crystals. The NaCl crystal has 6 slip systems.

In a polycrystalline material, slip in any crystal has to be accommodated by slip in neighbouring crystals, if the grain boundaries are to remain continuous during slip. According to the Von Mises criterion, a minimum of five independent slip systems is necessary to maintain the integrity of the grain boundaries during plastic deformation. In FCC and BCC crystals, this condition is fulfilled. In HCP crystals, for slip on the basal close packed plane, there are only two independent slip systems. These correspond to any two of the three close packed directions. A slip displacement in the third direction can be expressed as the resultant of the slip along the other two directions and is therefore not independent. In polycrystalline form, HCP materials can deform only by slip on less common slip systems or alternatively by twinning.

The stress at which slip starts in a crystal depends on the relative orientation of the stress axis with respect to the slip plane and the slip direction. When a tensile stress σ is applied to a crystal, as illustrated in Fig. 11.3, the shear stress τ

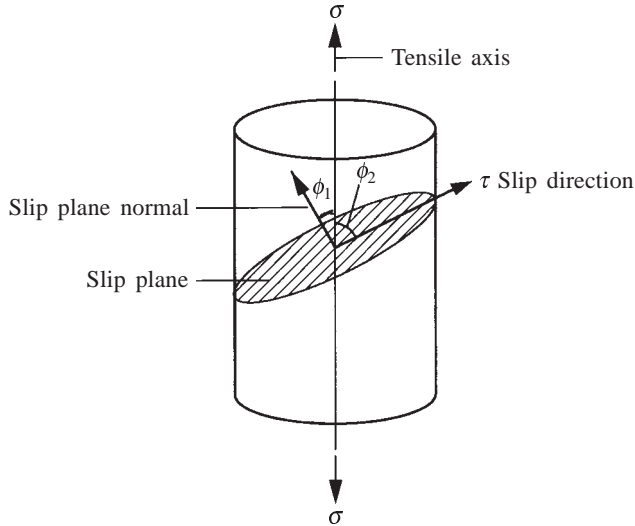


Fig. 11.3 The resolved shear stress τ on a slip plane along a slip direction depends on their orientation with respect to the tensile axis.

resolved on a slip plane whose normal makes an angle of ϕ_1 with the stress axis, along a slip direction inclined at an angle of ϕ_2 to the stress axis, is given by

$$\tau = \sigma \cos \phi_1 \cos \phi_2 \quad (11.7)$$

This resolved shear stress should reach a critical value called the *critical resolved shear stress* (CRSS) for plastic deformation to start. It is evident that all slip systems in a crystal will not have the same resolved shear stress for a given tensile stress along an axis. As the applied tensile stress is increased from zero, deformation will be initiated *first* on that slip system for which *the resolved shear stress is a maximum* and so reaches the critical value first.

Example 11.1 The axis of a cylindrical crystal of copper lies in the (111) plane. Show that, when a tensile stress σ is applied along the axis of the crystal, the resolved shear stress τ in any direction on this plane is zero. Does this mean that this crystal will not plastically deform however large is the applied stress?

Solution The angle ϕ_1 is 90° for the plane (111) so that $\cos \phi_1 = 0$. So, from Eq. (11.7), the resolved shear stress τ for any value of ϕ_2 is zero. This does not mean that the crystal will not plastically deform, as all planes in the family of {111} are not parallel to one another. There are planes other than (111), which will have a finite resolved shear stress. As soon as this reaches the critical value, plastic deformation will be initiated.

11.3 The Shear Strength of Perfect and Real Crystals

Figure 11.2b illustrates how a perfect crystal can undergo plastic deformation by the slip process. By using simplifying assumptions, we can estimate the stress required to cause slip in the perfect crystal. Figure 11.4 shows the stress-potential

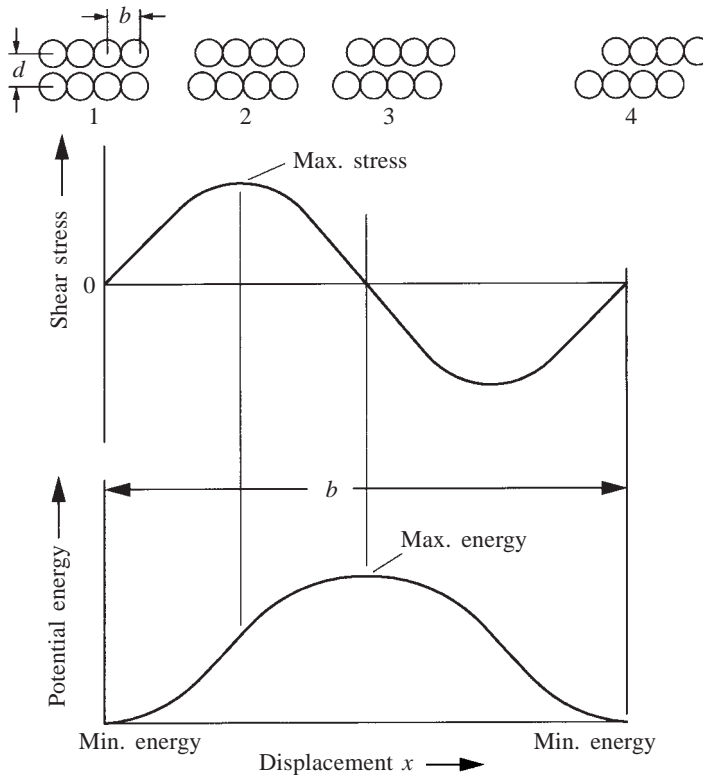


Fig. 11.4 Stress-potential energy-displacement curves during shear over one interatomic distance in a perfect crystal.

energy-displacement relationships, when a plane perpendicular to the paper (denoted by a row of atoms in the figure) is sheared towards the right by one interatomic distance with respect to another parallel plane (denoted by a second row of atoms). At position 1, no stress is applied and the potential energy is a minimum. The stress τ required to cause a displacement x will be proportional to dW/dx , where W is the potential energy. As the atoms of the top plane are displaced to position 2, the applied stress reaches its maximum value, corresponding to the inflection point on the potential energy curve. Position 2 is quarter way from position 1 in going to the next equilibrium position. Beyond position 2, the slope of the potential energy-displacement curve decreases. Consequently, a progressively-decreasing amount of stress will be required to cause further displacement, until the stress falls to zero at the half way position 3.

This position is in *unstable equilibrium* (recall Sec. 2.1) and the potential energy here is a maximum. Beyond 3, the slope of the potential energy curve becomes negative. The stress is also negative till the next equilibrium position is reached. The negative sign of the stress means that a stress in the opposite direction is necessary to stop the top plane from sliding into position 4, after it has crossed position 3. If the stress-displacement curve is taken to be sinusoidal, we can write:

$$\tau = K \sin \left(\frac{2\pi x}{b} \right) \quad (11.8)$$

where b is the atomic spacing in the direction of displacement and K is a constant. For small displacements, $2\pi x/b \ll 1$, and so we can write:

$$\tau = K \left(\frac{2\pi x}{b} \right) \quad (11.9)$$

Also, for small displacements, Hooke's law is valid and hence

$$\tau = \mu \left(\frac{x}{d} \right) \quad (11.10)$$

where μ is the shear modulus of the crystal and x/d is the shear strain, d being the distance of separation between the planes. Combining Eqs. (11.9) and (11.10), we obtain

$$K = \frac{\mu b}{2\pi d} \quad (11.11)$$

Substituting for K in Eq. (11.8), we get

$$\tau = \frac{\mu b}{2\pi d} \sin \left(\frac{2\pi x}{d} \right) \quad (11.12)$$

The stress τ_{\max} that is necessary for causing plastic deformation in a perfect crystal corresponds to the maximum amplitude of the sinusoidal wave and is equal to $\mu b/(2\pi d)$; b/d is about 1 for many crystals so that $\tau_{\max} \sim \mu/6$.

The shear modulus for typical crystals that exhibit plastic deformation is in the range 10–100 GN m⁻², so that the maximum stress τ_{\max} should be in the range 2–20 GN m⁻² or 2000–20 000 MN m⁻². The experimental values of the critical resolved shear stress (CRSS) for initiating plastic deformation in some typical crystals are given in Table 11.2. These values are very much lower than the range 2000–20 000 MN m⁻². Table 11.2 also gives the shear moduli and the ratio μ/τ for the same crystals. The ratio μ/τ is predicted to be around 6 on the basis of the perfect crystal calculations above. But the observed ratios given in Table 11.2 fall in the range 5000–100 000, indicating again the large discrepancy between the above model of plastic deformation and the experimental results. *Real crystals deform at a much lower stress than predicted by the model for a perfect crystal.* This discrepancy can be explained only on the basis of the *presence of dislocations* in real crystals. *The measured CRSS values should be*

TABLE 11.2
Observed Shear Strengths and Shear Moduli of Some Crystals

Crystal	Structure	CRSS τ , MN m ⁻²	μ , GN m ⁻²	μ/τ
Copper	FCC	0.5	44	88 000
Aluminium	FCC	0.75	25	33 000
Gold	FCC	0.5	26	52 000
Nickel	FCC	5	70	14 000
Silver	FCC	0.5	26	52 000
Iron	BCC	15	70	4 700
Zinc	HCP	0.3	33	110 000

associated with the stress required to move a dislocation that is already present in a real crystal and not with the stress required to shear a perfect crystal.

When a shear stress is applied parallel to a slip plane in a crystal, the dislocations on it move. As they sweep through the crystal, *each one of them causes a displacement* of the top part of the crystal with respect to the bottom by a distance *equal to the magnitude of its Burgers vector*. Unlike in the perfect crystal model of plastic deformation, the shear does not occur by the simultaneous displacement of all the atoms in a slip plane. The dislocation moves in a *sequential fashion*, occupying successive positions during its motion. Recall the atomic displacements during the motion of an edge dislocation shown in Fig. 6.12. During each step of the motion, the displacement of the atoms surrounding the dislocation is only a *fraction of an interatomic distance* and not a full one, as in the case of the perfect crystal. The incomplete plane of atoms is not bodily shifted but the configuration of the dislocation moves by *small adjustments in the bond lengths* in the dislocation region.

The analogy of the motion of a caterpillar helps in visualizing the dislocation motion. As illustrated in Fig. 11.5, the caterpillar moves by making a hump on its back and then transferring the hump from the tail to the head. It



Fig. 11.5 The motion of a caterpillar.

does not move forward all along its length simultaneously, but the hump moves sequentially. The frictional forces between the hump and the ground are absent, just as the bonds in the dislocation region are strained, though not broken. Therefore, the caterpillar spends much less effort in moving the hump. The magnitude of the displacement caused by the hump moving from one end to the other depends on the size of the hump. This aspect bears a direct analogy to the magnitude of the displacement caused by a moving dislocation, which is equal to its Burgers vector.

Normally, dislocations are always present in crystals. *Whiskers* are special crystals, which are very thin and almost free of dislocations. Such crystals can withstand stresses much higher than ordinary crystals, without undergoing plastic deformation. If, however, a dislocation is introduced accidentally, for example at the surface, the crystal abruptly loses all its strength and there is a big drop in the stress required to cause further strain, which is permanent. This is illustrated for a copper whisker in Fig. 11.6. The maximum stress the whisker withstands is

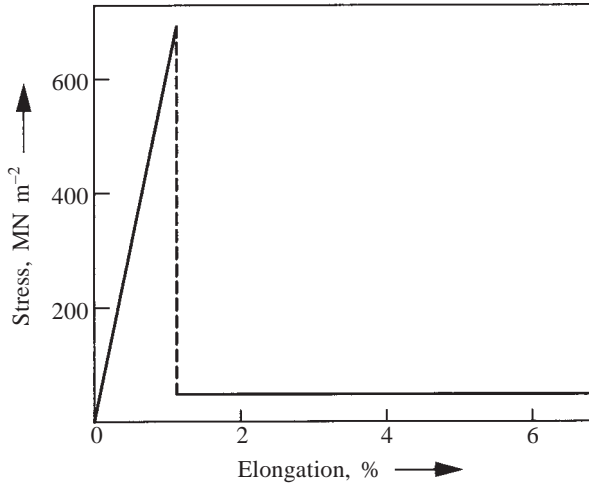


Fig. 11.6 A copper whisker deforms plastically, as soon as a dislocation is created, with a big drop in the stress required to cause further strain.

700 MN m^{-2} , but on the introduction of a dislocation, this stress falls precipitously to a much lower value and considerable plastic deformation ensues. (Compare the maximum stress that the copper whisker can carry, with the CRSS of an ordinary copper crystal listed in Table 11.2.)

11.4 The Stress to Move a Dislocation

The stress required to move a dislocation in a crystal in the absence of other imperfections and impurities has been computed by R. Peierls and F.R.N. Nabarro. These calculations from first principles are not accurate enough to predict the CRSS of different crystals and to correlate it with the observed value. However, we are in a position to understand the differences in the plastic deformation behaviour of crystals in a qualitative way.

Consider the two edge dislocations sketched in Fig. 11.7. In Fig. 11.7a, the dislocation is *stiff*, that is, no relaxing displacements have taken place in the adjacent planes surrounding the dislocation region. In Fig. 11.7b, such displacements have occurred around the incomplete plane. In the first case, the

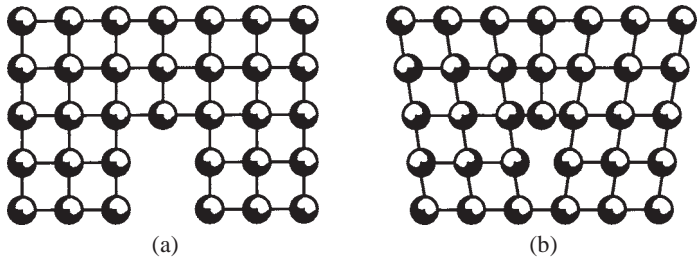


Fig. 11.7 An unrelaxed and a relaxed edge dislocation.

bond lengths are the normal value around the dislocation except in the region below the incomplete plane, where the bonds are virtually broken. In the second case, the strains are distributed more evenly among the bonds around the edge, resulting in some compression above and some tension below. Here, we say that the dislocation is *relaxed*. The width of a dislocation is a measure of the distance on either side of the dislocation, up to which the stress-relaxing displacements are appreciable. When a dislocation is said to be wide, it is meant that the displacements are distributed over an appreciable distance. When a wide dislocation moves, the adjustments that take place in the bond lengths in the dislocation region are distributed over a number of bonds in the region, but *the change in anyone bond length is very small*. In a narrow dislocation such as that shown in Fig. 11.7a, it is clear that, during the motion of the dislocation, the row of atoms below the slip plane has to move through a distance of one full interatomic distance. Therefore, *narrow dislocations are more difficult to move than wide dislocations*.

Peierls and Nabarro related the stress to move a dislocation τ_{PN} to its width w

$$\tau_{PN} = \mu \exp \left(- \frac{2\pi w}{b} \right) \tag{11.13}$$

where μ is the shear modulus of the crystal and b is the magnitude of the Burgers vector of the dislocation. Because of the exponential form of this relationship, the Peierls–Nabarro stress (τ_{PN}) is a very sensitive function of the width and the Burgers vector. The effect of width on τ_{PN} is illustrated as follows:

Width w	0	b	$5b$	$10b$
τ_{PN}	μ	$\mu/400$	$\mu/10^{14}$	$\mu/10^{27}$

When the dislocation has zero width, the crystal requires a very high stress to move a dislocation, comparable to the stress to deform a perfect crystal. Even as the width increases to $10b$, the stress falls to negligible values. A very precise determination of the dislocation width is then necessary to calculate the P–N stress in any reliable manner. Unfortunately, such measurements are inherently difficult and have not been made. We have to limit our discussion to the qualitative variation of the P–N stress.

The nature of the chemical bonding in a crystal determines the extent of the relaxation and the width of a dislocation. In *covalent crystals*, where the bonding

is strong and directional, the relaxation is small resulting in narrow dislocations. The Peierls–Nabarro stress is correspondingly high. The application of a tensile stress usually results in *brittle fracture* by crack propagation, before the stress required to move a dislocation is attained. Hence, covalent crystals such as diamond and silicon are brittle. They do not undergo plastic deformation. In *typical metallic crystals*, such as copper, the bonds are nondirectional and not so strong as in a covalent crystal. So, the dislocations are wide and the Peierls–Nabarro stress is *low*. These crystals exhibit a considerable amount of plastic deformation and are said to be ductile. A copper wire can be cold drawn to hundred times its original length without breaking, in spite of the work hardening that occurs during drawing. In metals of the *transition groups*, such as iron, some covalent character persists due to *d* orbital bonding which is directional and, correspondingly, the transition metal crystals are harder than copper. They cannot be cold worked to the same extent as copper.

In *ionic crystals*, the bonds are of moderate strength and nondirectional. From this, one would have expected simple ionic crystals to have some ability to undergo plastic deformation. However, plastic deformation occurs here only under special circumstances when the crystal surface is free of cracks that can cause brittle failure. The main reason for this behaviour seems to lie in the fact that the Burgers vector of dislocations in ionic crystals tends to be large. *b* is in the denominator of the exponential in Eq. (11.13). This tends to increase the Peierls–Nabarro stress. Only dislocations with a full lattice translation for the Burgers vector will not bring two cations as nearest neighbours during the slip process. The Burgers vector of a full dislocation in NaCl is 3.95 Å, as compared to that in copper, which is 2.55 Å.

Intermetallic compounds and other complex crystal structures such as Fe₃C and CuAl₂ do not have favourable crystal planes and directions for easy slip and therefore tend to be brittle. Ordered compounds such as CuZn require dislocations to move in pairs in order to preserve the order during slip. Hence, they possess very limited ductility.

11.5 The Effect of Temperature on the Stress to Move a Dislocation

The motion of a dislocation can be assisted by thermal energy. The rate of plastic deformation, that is, the strain rate ($d\varepsilon/dt$) is proportional to a Boltzmann's probability factor:

$$\frac{d\varepsilon}{dt} \propto \exp\left(-\frac{Q}{kT}\right) \quad (11.14)$$

where *Q* is the activation energy for dislocation motion. *Q* arises from the activation barrier that a dislocation faces in moving from one minimum energy position to the next. This barrier will be much smaller than the barrier shown for shear of a perfect crystal in Fig. 11.4. The strain rate during plastic deformation can be expressed in terms of the velocity v_d of the dislocations:

$$\frac{d\varepsilon}{dt} = \rho b v_d \quad (11.15)$$

where ρ is the density of mobile dislocations and b is their Burgers vector.

Equation (11.14) indicates that a dislocation can change its position solely due to thermal fluctuations even without the application of an external stress. Such motion will be random, where a dislocation line will have equal probability of moving to the next minimum energy position on either side of the line. Therefore, such motion does not result in plastic deformation. A certain threshold stress is necessary for dislocations to move in a specified direction. In addition to this, the applied stress can lower the activation energy Q . Let τ_{PN} be the stress to move a dislocation in the absence of thermal energy, that is, at 0 K. (Note that the Peierls–Nabarro equation (11.13) does not have a kT term.) At 0 K, an applied stress τ_{app} equal to τ_{PN} is necessary to move a dislocation, as no assistance from thermal energy is available. At higher temperatures, an applied stress lower than the P–N stress can move a dislocation, using thermal energy. The two stress terms can be multiplied by a volume term called the *activation volume* v to relate them to Q :

$$Q = (\tau_{\text{PN}} - \tau_{\text{app}}) v \quad (11.16)$$

Substituting Eq. (11.16) in Eq. (11.14), we have

$$\frac{d\varepsilon}{dt} \propto \exp \left[-\frac{(\tau_{\text{PN}} - \tau_{\text{app}}) v}{kT} \right] \quad (11.17)$$

This equation tells us that, for a constant strain rate, the term within the exponential is constant. τ_{PN} as given by Eq. (11.13) is a constant for a particular crystal. For our purposes, we can assume the activation volume v to be constant. It follows then that, with increasing temperature, less and less applied stress is required to cause plastic deformation at a specified strain rate. Alternatively, to cause plastic deformation at an increasing strain rate at a constant temperature, the applied stress should be increased. By taking logarithms, Eq. (11.17) can be rewritten as

$$\ln \left(\frac{d\varepsilon}{dt} \right) = -\frac{(\tau_{\text{PN}} - \tau_{\text{app}}) v}{kT} + A' = A + B \tau_{\text{app}} \quad (11.18)$$

where A , A' and B are constants at constant temperature.

Example 11.2 The activation volume for dislocation motion in a crystal is $20 b^3$, where b is the Burgers vector of the moving dislocation. $b = 2 \text{ \AA}$. The P–N stress for this crystal is 1000 MN m^{-2} . For a specified rate of dislocation motion, the activation energy $Q = 40 kT$. Calculate the stress required to move the dislocation at (i) 0 K, (ii) 100 K, (iii) 300 K, and (iv) 500 K.

Solution (i) At 0 K, $kT = 0$. The stress required to move the dislocation is equal to the P–N stress, 1000 MN m^{-2} .

(ii) At 100 K,

$$(\tau_{\text{PN}} - \tau_{\text{app}})20b^3 = Q = 40kT = 5.52 \times 10^{-20} \text{ J}$$

$$\tau_{\text{app}} = 1000 - (5.52 \times 10^{-20} \times 10^{-6}) / (20 \times 2^3 \times 10^{-30}) \text{ MN m}^{-2}$$

= 655 MN m⁻²—which is about two-thirds of the P–N stress.

(iii) At 300 K, the thermal energy kT is three times larger than at 100 K. Hence, the stress to be applied externally falls to zero. That is, the barrier to dislocation motion is entirely overcome by thermal fluctuations. However, a threshold stress to move the dislocations in a specified direction will be necessary.

(iv) The thermal energy available at 500 K is more than that needed to overcome the barrier to dislocation motion.

The activation volume denotes the volume over which the thermal energy is to concentrate in order to achieve the activation. In the above example, the activation volume was relatively small, $20b^3$. Dislocations face different kinds of obstacles during their motion through the crystal. In the Peierls–Nabarro model, the obstacle is due to lattice friction, as the barrier depends on the basic crystal structure. For other kinds of obstacles, the activation volume may be large. In such cases, the thermal energy may not be of much help in overcoming the obstacles. These obstacles are called *athermal obstacles* to dislocation motion. For example, the stress fields of the other dislocations in a crystal may pose an obstacle to the motion of a dislocation. Such obstacles have a large activation volume ($\sim 10\,000b^3$). Overcoming them would require virtually the entire stress to be applied externally, irrespective of the available thermal energy.

The yield stress of a number of crystalline materials is shown in Fig. 11.8 as a function of temperature. The P–N stress for a ductile crystal such as copper is

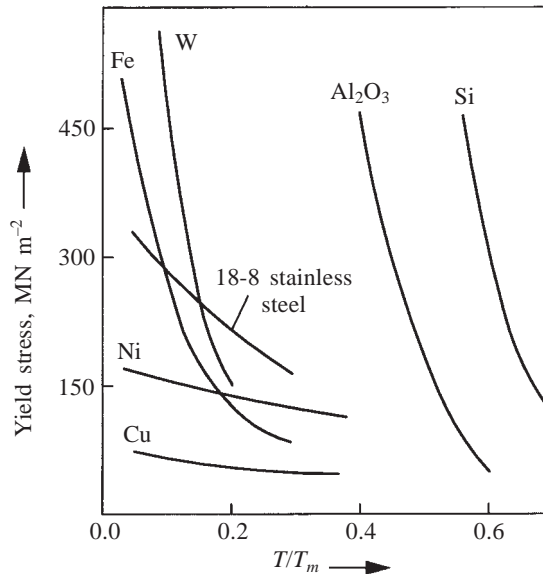


Fig. 11.8 The yield stress of some crystalline materials as a function of temperature.

so low that the thermal energy at most temperatures even below room temperature is sufficient to overcome the P–N stress. The temperature dependence of the yield stress is small, as only the threshold stress and the stress required to overcome athermal obstacles, if any, are needed at most temperatures. In a transition metal such as iron, the P–N barrier is comparatively large and a high stress is needed to move the dislocation near absolute zero. This stress decreases more or less linearly with increasing temperature. Eventually, around room temperature, the stress levels off to a nearly constant value, that is characteristic of the threshold stress and other athermal obstacles in the crystal. In covalent crystals such as diamond and silicon, the P–N stress is very high so that even at room temperature, a high stress is needed to move a dislocation. These crystals show no plastic deformation at room temperature. However, as the stress to move a dislocation is decreasing with increasing temperature, these crystals may exhibit some plastic deformation at elevated temperatures.

11.6 Multiplication of Dislocations during Deformation

In an annealed crystal, the dislocation density is about 10^{10} m^{-2} . On application of a stress, if all of these were to move out and disappear at the surface of the crystal, we would be left with a nearly perfect crystal with very few dislocations in it. Experimental observations are quite contrary to this picture. During plastic deformation, the dislocation density in the crystal *increases* by two to six orders of magnitude, depending on the amount of deformation undergone by the crystal. A lightly cold worked crystal may have a dislocation density of 10^{12} m^{-2} , while a very heavily cold worked crystal may have 10^{16} m^{-2} of dislocations. This indicates that there are sources within the crystal, which generate new dislocations during plastic deformation. One such source is called the *Frank–Read source*.

In Fig. 11.9, a segment AB of a dislocation line lies on a slip plane that is the most common one for this crystal. A and B are dislocation nodes or points beyond which the dislocation line does not lie on the slip plane. Only the segment AB is capable of moving easily in response to an applied stress. Let a shear stress τ be applied parallel to the slip plane on which AB is lying. The segment cannot pull itself away from the anchoring points, because of the geometrical requirement that a dislocation cannot end abruptly within the crystal. It responds to the applied stress, by bending itself to position 2 about the fixed points A and B and thereby increasing its length. The total line energy of the segment increases in the process. The extra energy comes from the work done by the externally applied stress in bending the dislocation. Between positions 1 and 2, the slip displacement equal in magnitude to the Burgers vector \mathbf{b} of the dislocation occurs over the shaded area of the slip plane, see Fig. 11.9.

As the stress is increased, the dislocation bends progressively to a semicircular shape as illustrated. The \mathbf{t} vector indicated in Fig. 11.9 follows the bent dislocation line, going from B to A . On further increase of the stress, the dislocation *bends back on itself* and the slip displacement spreads to regions on

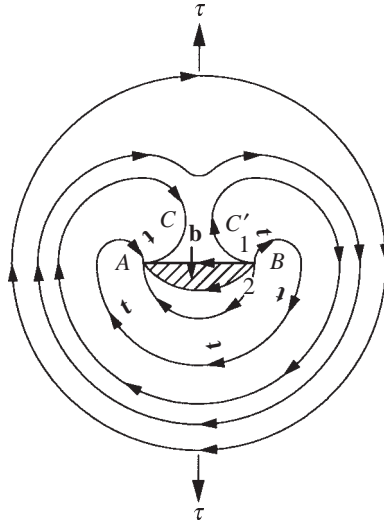


Fig. 11.9 Successive stages in the operation of a Frank-Read source.

the other side of the line AB . In one of these configurations, the direction of the dislocation line at point C is opposite in sense to the direction at point C' . The \mathbf{t} vectors are in opposite directions at these two points. As the Burgers vector is invariant, this means that the segment of the dislocation line near C is opposite in sign to that near C' . The two segments attract and *annihilate each other*. This produces a full dislocation loop and a left-over piece, which springs back to the initial position AB . Note that, during this entire cycle, the direction of dislocation motion at any point on the line is always perpendicular to the line at that point (except at A and B).

The cycle can now be repeated to produce another full loop and so on. A Frank-Read source can operate continuously and produce an indefinite number of loops, provided the loops produced move out and disappear at the surface of the crystal. If, however, the loops are piled up against an obstacle such as a grain boundary, back stress will begin to build up at the source, the operation of the source eventually coming to a halt.

The stress required to operate a Frank-Read source is given by

$$\tau = \frac{\mu b}{l} \quad (11.19)$$

where μ is the shear modulus of the crystal and l is the length of the segment AB . The longer is the segment, the lower is the stress required to operate it.

11.7 Work Hardening and Dynamic Recovery

In HCP crystals, the close packed basal planes are the common slip planes. They are all parallel to one another. In such cases, in a single crystal that is favourably

oriented with respect to the tensile axis, the Frank–Read sources in the basal planes can operate continuously for an unlimited period, as the loops produced by the sources can move out and disappear at the surface of the crystal. There is very little work hardening here, as the same source operates at the same resolved shear stress. The stress to cause plastic deformation remains virtually independent of the plastic strain. If, however, the tensile stress axis happens to be such that there is only a small amount of resolved shear stress on the basal planes, then slip on less common slip planes can also occur. These less common slip planes (prismatic and pyramidal planes) are nonparallel to the basal planes. Dislocations moving on nonparallel slip planes can intersect each other. In BCC and FCC crystals, the intersection of dislocations moving on nonparallel slip planes is very common. This intersection generally increases the stress required to cause plastic deformation with increasing amount of plastic strain, resulting in *work hardening*.

For example, consider the dislocations moving on nonparallel planes of an FCC crystal. The slip planes in FCC belong to the $\{111\}$ family. Let a dislocation of Burgers vector $\frac{1}{2}[0\bar{1}1]$ moving on $(\bar{1}11)$ plane meet another dislocation of Burgers vector $\frac{1}{2}[10\bar{1}]$ moving on $(1\bar{1}1)$ plane along the line of intersection of these two planes, Fig. 11.10. These two dislocations can interact and produce a third dislocation with a decrease in the potential energy:

$$\frac{1}{2}[0\bar{1}1] + \frac{1}{2}[10\bar{1}] \rightarrow \frac{1}{2}[1\bar{1}0] \quad (11.20)$$

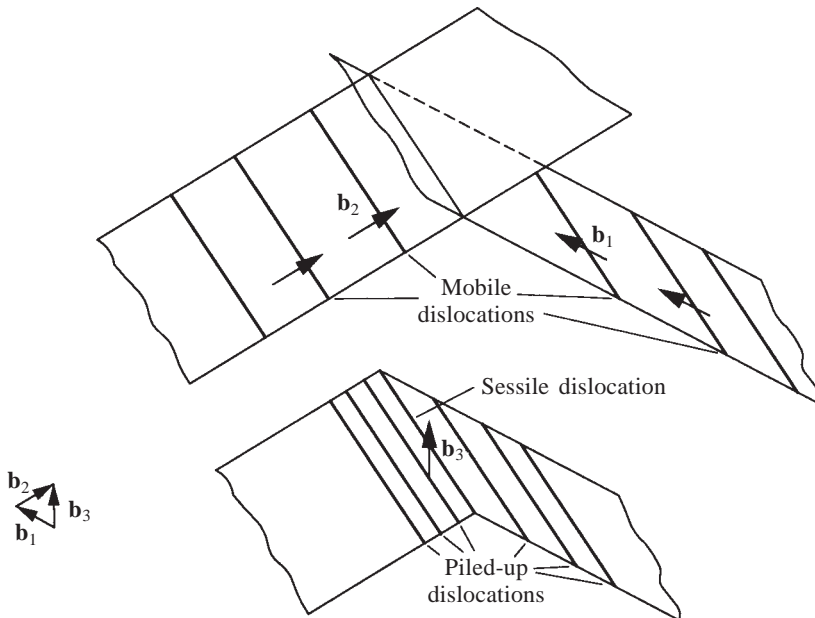


Fig. 11.10 Formation of a sessile dislocation in an FCC crystal by a dislocation reaction.

The strain energy decreases by 50% during this dislocation reaction. As the Burgers vector of the product dislocation lies in neither of these two slip planes, it becomes *immobile* (or sessile). Once the immobile dislocation forms by the dislocation reaction given above, it acts as an obstacle to the oncoming dislocations from the respective sources, which get piled up against the obstacle. A back stress builds up and the dislocation sources stop operating.

The *most effective* Frank–Read source (which is the longest) is expected to operate *first* (see Eq. (11.19)). Once this gets immobilized, less effective sources, consisting of shorter segments, have to operate. This requires an increase in the stress as given by Eq. (11.19). These sources will in turn get immobilized by the formation of the sessile dislocation. The plastic deformation can continue only with the activation of still shorter sources. Thus, a progressively increasing amount of stress is required to continue the plastic deformation process, resulting in *work hardening* even in single crystals. The formation of the immobile dislocation is at best one possible explanation for work hardening. There are several other detailed theories to explain work hardening, but we need not go into these.

Empirically, it is found that the shear stress τ to move a dislocation increases with increasing dislocation density ρ according to the equation

$$\tau = \tau_0 + A\sqrt{\rho} \quad (11.21)$$

where τ_0 is a base stress to move the dislocation in the crystal in the absence of other dislocations and A is a constant. This equation describes equally well the work hardening behaviour, as the dislocation density increases with increasing plastic strain.

For a soft crystal, the CRSS for initiation of plastic deformation is typically 0.5 MN m^{-2} . If the constant $A = 10 \text{ N m}^{-1}$, a low dislocation density of 10^{10} m^{-2} corresponding to an annealed crystal gives a CRSS of 1.5 MN m^{-2} . If the crystal is now heavily cold worked, the dislocation density can increase to 10^{14} m^{-2} and the flow stress τ will now be 100 MN m^{-2} , a large increase.

The stress-strain curves for specimens of mild steel with increasing amount of cold work are shown in Fig. 11.11. Cold working increases the yield stress

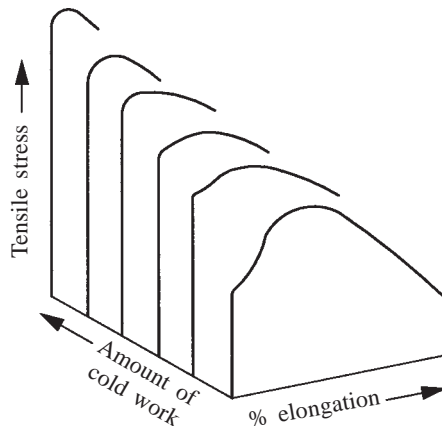


Fig. 11.11 The stress-strain curves for specimens of mild steel cold worked to different degrees.

but decreases the per cent elongation. Cold working is a common strengthening method used in engineering practice. Under suitable conditions, the dislocation density can be increased to very high values, with the strength approaching that of a perfect crystal.

Example 11.3 The length of a dislocation line between two pinning points is on an average equal to the reciprocal of the square root of the dislocation density in a crystal. Calculate the dislocation density in copper, work hardened to a stage where slip occurs at a shear stress of 35 MN m^{-2} .

Solution The shear modulus of copper is 44 GN m^{-2} . The Burgers vector of dislocations in copper is $3.61/\sqrt{2} = 2.55 \text{ \AA}$. From Eq. (11.19),

$$l = \frac{\mu b}{\tau} = \frac{44 \times 10^9 \times 2.55 \times 10^{-10}}{35 \times 10^6} = 0.32 \text{ \mu m}$$

$$\rho = \frac{1}{l^2} = \frac{1}{0.32^2 \times 10^{-12}} = 1.0 \times 10^{12} \text{ m}^{-2}$$

Single crystal experiments show that after a certain stress level is reached, the rate of strain hardening decreases with further strain. This phenomenon is called *dynamic recovery*. This stress level is sufficient to activate screw dislocations to *cross-slip*. As the Burgers vector of the screw dislocations is parallel to the dislocation line, they are free to move on any out of several equivalent slip planes. When deformation is initiated, they move on those planes where the resolved shear stress is a maximum. If they happen to pile up against an immobile dislocation, they can cross-slip provided the applied shear stress is sufficiently large to give the necessary resolved shear stress on the new slip plane.

11.8 The Effect of Grain Size on Dislocation Motion

Grain boundaries provide obstacles to dislocation motion. As the orientation of the crystals on either side of a grain boundary is different and random, a dislocation moving on a common slip plane in one crystal can rarely move onto a similar slip plane in the adjacent crystal. In addition, the crystals are separated by a thin noncrystalline region, which is the characteristic structure of a large angle grain boundary. Hence, dislocations are stopped by a grain boundary and pile up against it. The smaller is the grain size, the more frequent is the pile up of dislocations. A twin boundary can act as an obstacle to dislocations in a similar manner. Figure 11.12 shows the pile up of dislocations against twin boundaries in a Cu-8% Al alloy.

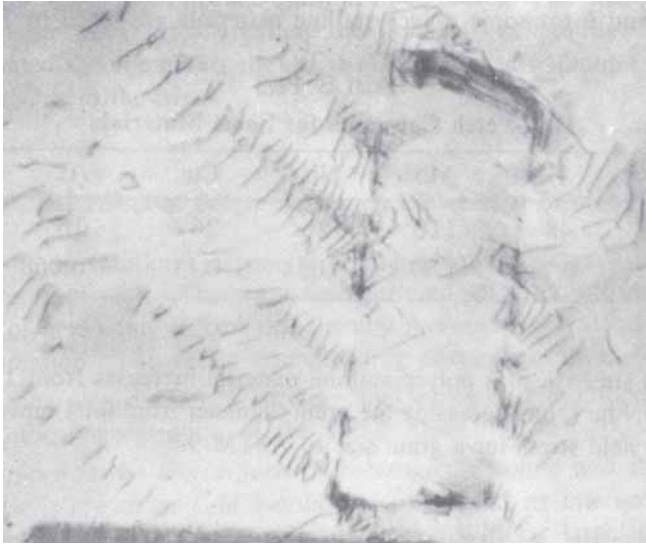


Fig. 11.12 Electron micrograph of dislocations piled up against twin boundaries in Cu-8%Al alloy. Magnification 15 000 \times . (J. Nutting and R.G. Baker, *The Microstructure of Metals*, by permission from the Institute of Metals, London.)

In a piled up array of stationary dislocations, the stress concentration at the leading dislocation (which is nearest to the grain boundary) is equal to the product of the resolved shear stress on the slip plane and the number of other dislocations in the pile-up. The magnitude of this concentration depends on the grain size which determines the average distance from a dislocation source to the grain boundary. In large grain sizes, the concentration factor may be as high as 100 and is more than sufficient to initiate plastic deformation in the adjoining grain, by activating the Frank–Read sources in it. The spreading of the plastic deformation from one grain to the next is then obviously a function of the grain size, and is more effective for a large grain size. The yield strength of a polycrystalline material as observed macroscopically in a tensile test requires the initiation of the plastic deformation in a large number of grains by the spreading process. Hence, the observed yield stress increases with decreasing grain size.

E.O. Hall and N.J. Petch have derived the following equation for the yield stress σ_y of a polycrystalline material:

$$\sigma_y = \sigma_i + kd^{-1/2} \quad (11.22)$$

where σ_i is the yield stress for a crystal of the same material where there are no grain boundaries, k is a constant and d is the average grain diameter. The yield strength increases with decreasing grain diameter.

The grain diameter can be approximately calculated from the ASTM specification for grain size. For an ASTM number n , the number of grains per square inch (per 645 mm²) at a magnification of 100 \times is equal to 2^{n-1} . Grain size number ASTM 1 corresponds to 1 grain per square inch at a magnification of 100 \times or $10^4/645 = 15.5$ grains per mm² without any magnification. This can

be approximated to a grain diameter of $1/\sqrt{15.5} = 0.25$ mm. The values of σ_i and k for some polycrystalline materials are listed in Table 11.3.

TABLE 11.3
Hall–Petch Constants for Some Materials

Material	Fe	Mo	Nb	Cu	Al	Zn
σ_i , MN m ⁻²	48	110	126	26	16	33
k , MN m ^{-3/2}	0.71	1.77	0.034	0.11	0.068	0.22

Example 11.4 The yield strength of a polycrystalline material increases from 120 MN m⁻² to 220 MN m⁻², on decreasing the grain diameter from 0.04 mm to 0.01 mm. Find the yield stress for a grain size of ASTM 9.

Solution By substituting the two yield stresses and the two grain diameters in Eq. (11.22), we obtain

$$\begin{aligned} 120 &= \sigma_i + k (0.04 \times 10^{-3})^{-1/2} \\ 220 &= \sigma_i + k (0.01 \times 10^{-3})^{-1/2} \end{aligned}$$

Solving for σ_i and k , we get

$$\begin{aligned} \sigma_i &= 20 \text{ MN m}^{-2} \\ k &= 0.633 \text{ MN m}^{-3/2}. \end{aligned}$$

Grain diameter for ASTM 9 is

$$d = \frac{1}{\sqrt{10^4 \times 256/645}} = 0.0159 \text{ mm}$$

Yield stress for this grain size is

$$\begin{aligned} \sigma_y &= 20 + 0.633 \times (0.0159 \times 10^{-3})^{-1/2} \\ &= 179 \text{ MN m}^{-2} \end{aligned}$$

The normal range of grain sizes in metals varies in ASTM specification from ASTM 1 to ASTM 8. In recent years, ultra fine grain sizes have been produced in *microalloyed steels*, which contain small quantities of strong carbide forming elements such as Nb, V or Ti. These elements go into solution in austenite on reheating the steel billet to about 1300°C. When the steel is subsequently hot rolled, the temperature falls gradually and the solubility of the alloying elements decreases. Very fine particles of alloy carbides precipitate from austenite. These precipitates effectively pin down the migrating grain boundaries during repeated recrystallization of the deformed austenite between passes at successive stands of the rolling mill. This decreased growth rate due to pinning yields a fine-grained recrystallized austenite. A fine-grained austenite provides more potential sites at the grain boundaries for the nucleation of ferrite as the temperature further drops.

The end result is thus a very fine grain size of ferrite in the steel of about 2–3 μm (ASTM 14–15). The yield strength increases as per the Hall–Petch equation. The increase is about 50%. The other advantages are that the excellent weldability of mild steel remains unaffected and the microalloyed steel is produced without a heat-treatment step, which adds to the cost.

11.9 The Effect of Solute Atoms on Dislocation Motion

Solid solutions, in general, offer a greater resistance to dislocation motion than pure crystals. The stress fields around solute atoms interact with the stress field of a moving dislocation, thereby increasing the stress required for plastic deformation. The solute strengthening effect depends mainly on two factors:

- (i) The size difference between the solute and the solvent atom
- (ii) The concentration of the solute atom.

The more the size difference between the solute and the solvent, the more intense is the stress field around the solute and its interaction with the moving dislocation is correspondingly stronger. With a large concentration of the solute, the moving dislocation interacts with the solute stress fields at many points along its length. The effect of different solutes on the yield stress of copper is shown in Fig. 11.13, as a function of the concentration of the solute. The yield stress is approximately proportional to the concentration of the solute. The slope

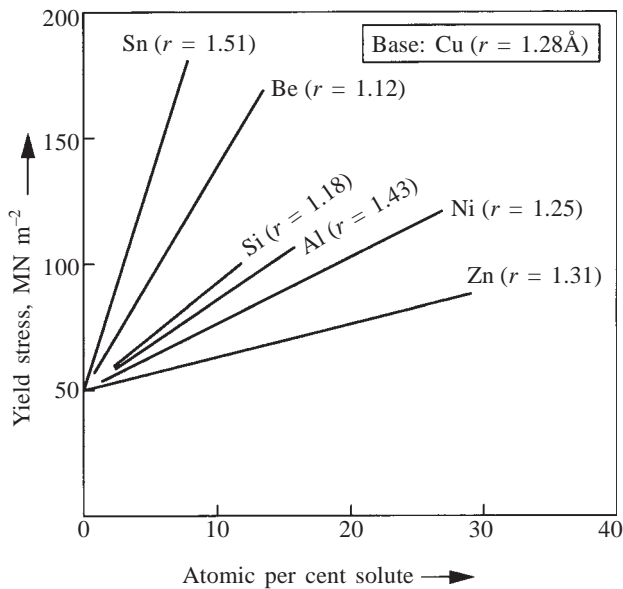


Fig. 11.13 Solute strengthening of a copper crystal by solutes of different sizes.

of the straight lines in the figure depends on the size difference. For example, zinc and nickel have a size difference of 0.03 \AA with respect to copper:

$$|r_{\text{Zn}} - r_{\text{Cu}}| = 0.03 \text{ \AA}$$

$$|r_{\text{Ni}} - r_{\text{Cu}}| = 0.03 \text{ \AA}$$

This size difference is small and, accordingly, the strengthening effect of zinc and nickel is the least among the solutes shown in Fig. 11.13. The size difference between copper and tin is large and the slope of the line for tin is also large.

For the same size difference, the smaller atom, viz. nickel, produces a greater strengthening effect than zinc. In addition to the size difference, the intensity of the stress field around a solute atom is also dependent on the elastic modulus of the solute. Nickel having a higher elastic modulus than zinc produces a more intense stress field and a greater strengthening effect.

In order to improve the strength of a crystal by solute strengthening, it would appear that the maximum size difference coupled with the maximum concentration of the solute would give the best results. However, these two factors are mutually exclusive. Recall from Hume–Ruthery's rules: the *more the size difference between the solute and the solvent, the smaller is the equilibrium solubility*. The solubility can be increased by producing a supersaturated metastable solid solution by quenching from an elevated temperature.

The effect of the size difference and the concentration of the solute described above for substitutional solutes is also valid for interstitial solutes. The interstitial atoms are usually larger than the interstitial voids they occupy. Here, the strengthening effect can be very strong, except that the equilibrium solubility tends to be small. Nature has provided mankind with a unique reaction in steel, which permits an unusually large solubility of interstitial carbon in iron in spite of a very unfavourable size effect. Martensite in steels is a supersaturated solution of carbon in iron, obtained by quenching the steel and not allowing the carbon to diffuse out of the iron lattice. For example, an eutectoid steel contains 0.8% carbon in martensite, which is some 40 times more than the equilibrium solubility in ferrite ($\sim 0.02\%$) at the eutectoid temperature. On quenching, the carbon gets trapped in the interstitial positions, as the austenite shears over to form the martensitic structure. The stress field produced by the oversized carbon atoms is so intense that the dislocation motion is very effectively hindered. Indeed, it is so effective that it becomes necessary to temper the martensite to restore some ductility, at the expense of some hardness.

In addition to solute strengthening, the martensitic plates may contain either a high dislocation density or very fine transformation twins, depending on the type of martensite obtained. These features also aid in increasing the strength of martensite.

The carbon atoms in the ferrite phase are responsible for the occurrence of a sharp yield point in mild steel. They segregate around the dislocation cores and reduce the total distortional energy. The segregation is known as a *Cottrell atmosphere* and the pinning effect of the atmosphere on the dislocations raises the yield stress of the crystal. The dislocations are strongly locked by the atmosphere and do not get freed for motion during plastic deformation. Very few free dislocations are available at the start of the plastic

deformation near the upper yield point of the mild steel. Once the deformation starts, the rapid multiplication of new dislocations and the stress dependence of their velocity are responsible for a fall in stress to the lower yield point at the beginning of plastic region. Thereafter, the stress-strain curve shows the usual work hardening characteristics.

After a mild steel specimen has been strained beyond the yield point, if the load is released and reapplied immediately, no yield point is observed, as carbon atoms have not had the time to diffuse to the new dislocation cores for forming atmospheres. If the steel is aged for some time at room temperature or above, the carbon atoms diffuse to the dislocations and lock them up. So, the yield point returns after ageing.

11.10 The Effect of Precipitate Particles on Dislocation Motion

Precipitate particles in a matrix provide obstacles to the motion of dislocations. The dislocations moving through the matrix phase have two alternatives. They can either *cut through the precipitate particles or bend around and bypass them*, Fig. 11.14. The first alternative will be possible only when the slip plane is continuous from the matrix through the precipitate particle and when the stress to move a dislocation in the crystal structure of the precipitate particle is comparable to that in the matrix. Cutting through occurs for very small precipitate particles, which can be considered to be groups of segregated solute atoms. Cutting through will not be possible, if there is an interface between the precipitate and the matrix and if the orientation changes abruptly at the interface. Under such circumstances, the dislocations bend around and bypass the precipitate particles, as illustrated in Figs. 11.14b and 11.14d. The mechanism of bypassing is very similar to the operation of a Frank–Read source. The stress required to bend a dislocation is inversely proportional to the average interparticle spacing, see Eq. (11.19). Effective strengthening is achieved in the bending process, when the precipitate particles are submicroscopic in size. The spacing between them should be typically a few hundred angstroms. Optimum ageing results in the right interparticle spacing. Overageing allows the fine particles to coalesce into larger and fewer particles. The interparticle spacing is increased thereby and the yield stress decreases.

Example 11.5 One million precipitate particles per unit volume of a precipitation hardened alloy coalesce into 1000 particles due to overageing. Assuming the particles to be uniformly distributed and their volume to be small compared to the volume of the matrix, estimate by how much the yield strength would have changed.

Solution When 10^6 particles coalesce to form 10^3 particles, the linear distance of separation between them increases to $\sqrt[3]{10^6/10^3} = 10$ times the initial value. The stress required to bend the dislocation around precipitate particle is inversely proportional to the spacing between the particles. So, the yield strength would have decreased to 10% of the initial value, due to the coarsening.

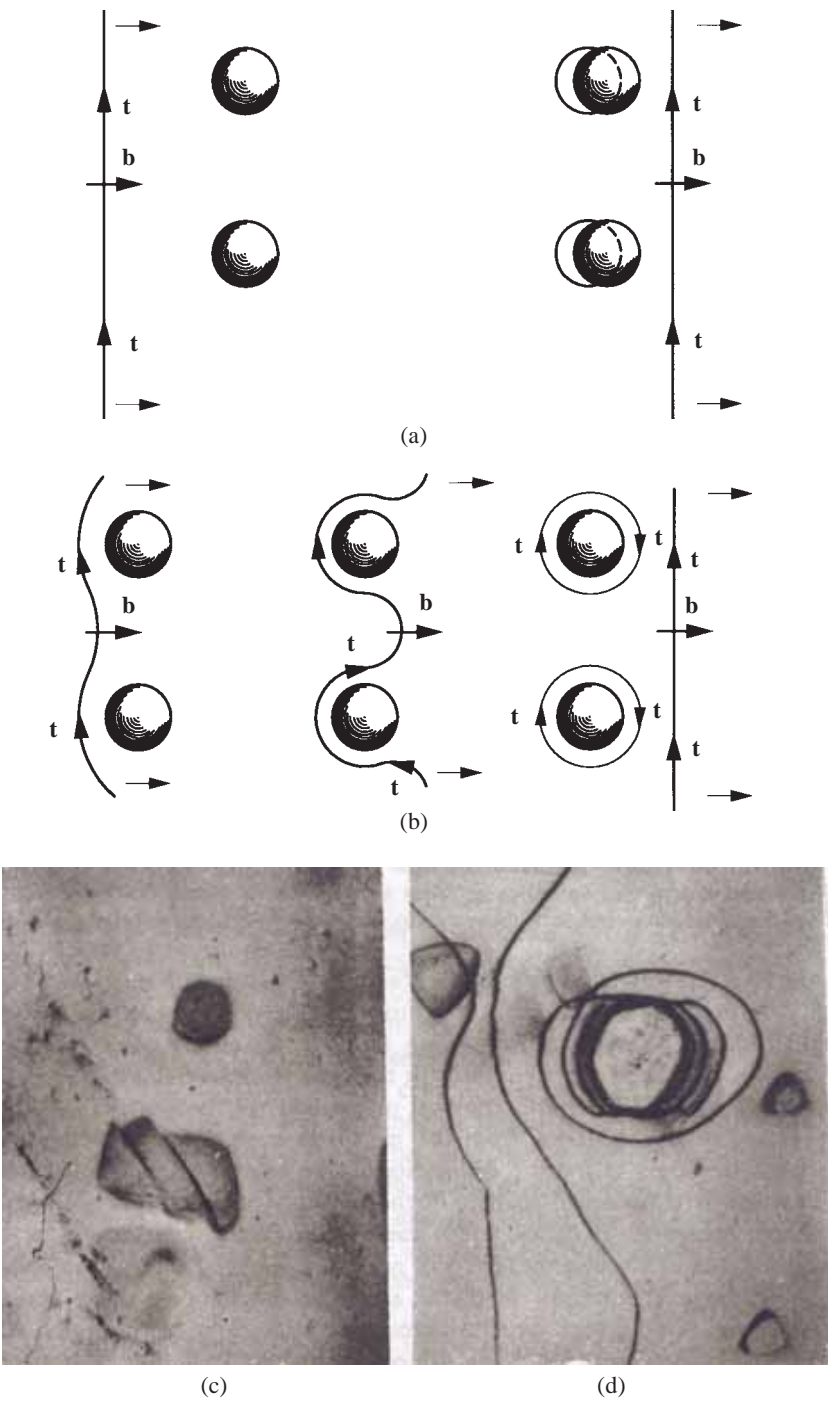


Fig. 11.14 A moving dislocation either (a) cuts through the precipitate particles or (b) bypasses them. Electron micrographs depicting these two processes are shown in (c) and (d). [(c) and (d) *Courtesy: F.J. Humphreys and V. Ramaswamy.*]

Table 11.4 lists the yield strength and per cent elongation of some precipitation hardened alloys.

TABLE 11.4
Yield Strength and Per cent Elongation of Some Precipitation Hardened Alloys

Alloy	Approximate composition	Heat-treated condition	Yield strength, MN m ⁻²	Per cent elongation
Al–Zn–Mg	5.6%Zn, 2.6%Mg	Annealed	103	16
		Aged	503	11
Cu–Be	2%Be	Annealed	221	47
		Aged	1069	7
Stainless steel	17%Cr, 7%Ni, 0.07% C	Annealed	324	39
		Aged	1276	9
Maraging steel	8%Co, 5%Mo, 0.8%Nb, 0.4%Ti	Annealed	648	18
		Aged	1530	11

11.11 Review of Strengthening Methods

As a review, the four main methods of strengthening crystalline materials against plastic yield are summarized below.

Real crystals are weak due to the presence of dislocations in them. Perfect crystals (whiskers) which are free of dislocations have very high strength, but they are so thin that they cannot be used as structural and machine components.

The key to increase the strength of crystalline materials against plastic yield is to increase the stress required to move a dislocation.

11.11.1 Strain Hardening

When metals are cold worked, their strength and hardness increase. Annealed crystals have a dislocation density of about 10^8 m^{-2} . This can be increased to 10^{10} – 10^{12} m^{-2} by moderate cold working and to 10^{14} – 10^{16} m^{-2} by heavy cold working. As the dislocation density increases, the stress required to move any one dislocation increases due to the interfering effect of the stress fields of the surrounding dislocations. This phenomenon is the basis of work hardening and is described by means of the following equation:

$$\tau = \tau_0 + A\sqrt{\rho}$$

where τ is the stress to move a dislocation in a matrix of dislocation density ρ , τ_0 is the stress to move the dislocation in the same matrix with zero dislocation density and A is a constant.

11.11.2 Grain Refinement

As the orientation changes at a grain boundary, the slip plane in a grain does not continue in the same direction beyond the boundary. So, dislocations gliding on a slip plane are unable to cross the boundary but get piled up against it. This happens more often in a fine grained material. The yield strength of a polycrystalline material is a function of its grain size as given by the Hall–Petch equation:

$$\sigma_y = \sigma_i + kd^{-1/2}$$

where σ_i is the yield strength at ‘infinite’ grain size (single crystal), k is the Hall–Petch constant, and d is the mean grain diameter, which can be derived from the ASTM grain size number.

The strengthening effect for a given grain size depends on the magnitude of the constant k . For example, k for BCC iron is $0.71 \text{ MN m}^{-3/2}$, whereas it is only 0.11 and 0.07 for the FCC metals, copper and aluminium. So, a given amount of grain refinement produces a greater strengthening effect in iron.

11.11.3 Solid Solution Strengthening

The strengthening effect of solute atoms in a crystal depends on the size difference between the solute and the solvent atoms and the concentration of the solute. The more is the difference in the atomic sizes, the more is the strengthening effect. Higher concentrations of the solute also increase the strength proportionately. It is not possible to dissolve a large concentration of a solute, which differs substantially in size from the solvent, as this would be contradicting the Hume–Rothery’s rules. However, this difficulty could be partly overcome by quenching from an elevated temperature and retaining the larger solubility at the higher temperature in the form of a supersaturated solution at the lower temperature. Quenching austenite retains all the carbon in solution in the product phase martensite, producing very hard steels.

11.11.4 Precipitation Strengthening

A microstructure that consists of a very fine, submicroscopic distribution of precipitates in a matrix has a good strength. The dislocations moving in the matrix are effectively hindered by the closely spaced precipitate particles, as they have to bend around and bypass the particles. The strengthening effect is inversely proportional to the particle spacing, the minimum attainable spacing being of the order of 100 \AA .

Such a microstructure tends to become unstable at elevated temperatures, as the particles undergo coarsening which increases the interparticle spacing and decreases the strength correspondingly. This happens during overageing in aluminium alloys.

CREEP

Creep is the permanent deformation of a material under load as a function of time. It is appreciable only at temperatures above $0.4T_m$. Room temperature for iron is $0.16T_m$ and for copper it is $0.22T_m$ so that creep at room temperature is negligible in these materials. On the other hand, room temperature is about half of the melting point (in kelvin) for lead and so it undergoes creep at room temperature. Creep of lead under its own weight is seen in old time roofs, where a thick rim of lead is found at the edge of the roofs.

11.12 Mechanisms of Creep

Figure 11.15 shows typical creep curves. The strain is plotted as a function of time under constant load or constant stress. The creep curve exhibits three

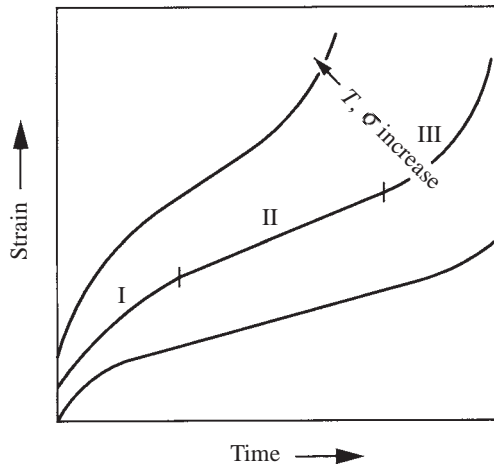


Fig. 11.15 Creep curves showing the increase in plastic strain as a function of time.

stages. In stage I, the creep rate decreases with time; the effect of work hardening is more than that of recovery in this stage. In stage II, the creep rate is a minimum and is constant with time; here, the work hardening and recovery processes are exactly balanced. In stage III, the creep rate increases with time until fracture occurs. In this stage, necking of the specimen starts and other processes that ultimately result in failure set in.

The temperature and time dependence of creep deformation indicates that it is a *thermally activated process*. Several atomic processes are known to be responsible for creep in crystalline materials, see Fig. 11.16.

When considering work hardening in Section 11.7, it was pointed out that dynamic recovery occurs in the later stages of the plastic deformation process. Dynamic recovery was attributed to the cross-slip of screw dislocations. In the low temperature region of creep, the *cross-slip* continues with the aid of thermal energy and causes further plastic strain as a function of time.

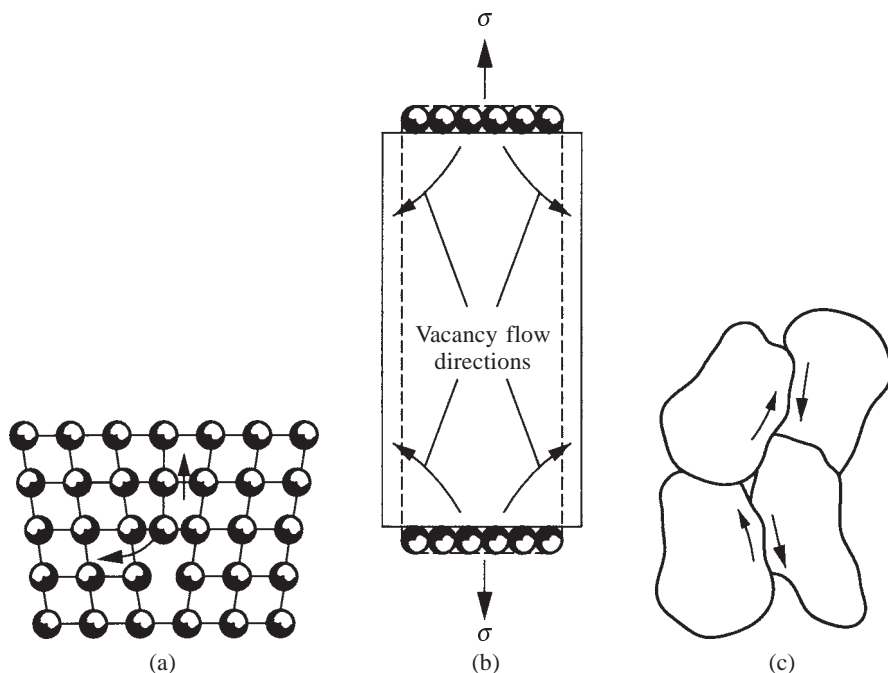


Fig. 11.16 Mechanisms of creep: (a) dislocation climb; (b) vacancy diffusion; and (c) grain boundary sliding.

The Burgers vector \mathbf{b} and the \mathbf{t} vector of edge dislocations are nonparallel. Therefore, they cannot cross-slip like screw dislocations. However, if the temperature is high enough for an appreciable diffusion rate of vacancies, these may diffuse to the edge dislocations or away from them, making them to *climb up or down*, see Fig. 11.16a. Thus edge dislocations piled up against an immobile dislocation can move to other parallel slip planes by climb and continue their motion in response to the stress. *The rate controlling step in the climb process is the diffusion of vacancies.* The measured activation energy for creep agrees with the activation energy for self-diffusion by the vacancy mechanism in a number of materials.

Another mechanism of creep is called *diffusional creep*. Here also the diffusion of vacancies controls the creep rate, but the mechanism does not involve the climb of edge dislocations. Vacancies move in response to the applied stress from surfaces of the specimen transverse to the stress axis to the surfaces that are parallel to the stress axis. Over a period of time, this movement would elongate the specimen in the direction of the stress axis and contract it in the transverse direction resulting in creep, as shown in Fig. 11.16b.

The third mechanism of creep is the *sliding of neighbouring grains* with respect to the boundary that separates them (see Fig. 11.16c). Grain boundaries lose their strength at a lower temperature than the grains themselves. This effect arises from the noncrystalline structure of the grain boundaries. As illustrated in

Fig. 9.17, the viscosity of a noncrystalline substance decreases continuously on heating into the liquid range, whereas a crystalline substance maintains its viscosity at a high value right up to the melting point. At temperatures above $0.5 T_m$, the viscosity of the grain boundaries is small enough for them to behave like a very viscous liquid separating the neighbouring grains and allowing them to slide against each other. This behaviour of the grain boundaries is in contrast to their low temperature behaviour. At low temperatures, they do not flow viscously, but provide effective obstacles to dislocation motion. *At high temperatures, the grain boundaries facilitate the deformation process by sliding, whereas at low temperatures, they increase the yield strength by stopping the dislocations.*

11.13 Creep Resistant Materials

Materials for machine parts and structural components used at elevated temperatures must be creep resistant. A number of engineering processes require the use of elevated temperatures. Cracking stills in the petroleum industry need high temperatures to accelerate the reaction rates (recall the Arrhenius law) and to control the reaction products. In energy conversion, the efficiency η of a heat cycle such as the Carnot cycle is dependent on the difference between the operating temperature T_2 and the sink temperature T_1 :

$$\eta = \frac{T_2 - T_1}{T_2} \quad (11.23)$$

A thermodynamic efficiency of less than 30% with the use of steam at atmospheric pressure can be increased to more than 75% for a gas turbine with an operating temperature of 900°C. Materials used for turbine blades must be creep resistant. They should not deform by creep during service. Creep can result in seizing of the blades with the turbine casing.

As a first requirement, the materials to be used at high temperatures must have a high melting point, as creep becomes significant at temperatures greater than $0.4T_m$. Refractory oxides such as MgO and Al₂O₃ have high melting points and, as such, are very suitable for high temperature use. Their brittleness limits their use to applications where only compressive stresses are encountered. Recent developments of tougher ceramics such as silicon nitride (Si₃N₄) indicate that selected parts of a heat engine such as piston rings and cylinder heads can be produced from ceramic materials.

Metals and alloys can be used under more versatile conditions. Most creep resistant alloys consist of a base metal of a fairly high melting point. Since the ability to fabricate into required shapes of the components, the relative cost of production and the density are the other considerations in the choice of a material, metals such as tungsten with a melting point of well over 3000°C have not found widespread use. The commonly used high temperature alloys are iron-base, nickel-base and cobalt-base alloys. All the three base metals have moderately high melting points around 1500°C. The creep resistance of these alloys in the temperature range from $0.5T_m$ up to the melting point is

considerably improved by a special strengthening process called *dispersion hardening*. In TD (thoria dispersed) nickel, fine particles of thoria (ThO_2) are dispersed in the nickel matrix and the interparticle distance is small enough for effective hindrance of dislocation motion in the matrix. TD nickel maintains its strength up to $0.9T_m$. The strengthening mechanism is similar to that of precipitation hardening discussed in Sec. 11.10 under plastic deformation. However, there is an important difference. At high temperatures, a precipitation-hardened alloy would lose its strength by the coarsening process. The tiny precipitate particles would coalesce into fewer and bigger particles, thereby increasing the interparticle distance and lowering the strength. The high surface area to volume ratio of the fine particles provides the driving force for the coarsening process. In dispersion hardening, on the other hand, oxide particles are embedded in a metallic matrix, with very little solubility of the oxide in the metallic phase. *This negligible solubility effectively prevents coarsening*, as the coalescence of the particles can occur only by the smaller particles dissolving and reprecipitating on larger particles. In nickel base superalloys, the coarsening is prevented by a different mechanism. Here, the precipitate particles of $\text{Ni}_3(\text{Ti}, \text{Al})$ form *an interface with the matrix of a very low energy*, about 0.005 J m^{-2} (5 erg/cm^2). As the decrease in the total surface energy is the driving force for coarsening, very little driving force is available here.

Among the other strengthening processes discussed under plastic deformation, cold working cannot be used for creep resistance. At temperatures above $0.4T_m$, recrystallization will occur quite readily and the cold-worked strength will be lost on recrystallization, recall Fig. 9.19. Solid solution strengthening can be used for better creep resistance, in the same way as in plastic deformation.

A fine grained material is desirable for better mechanical properties in a low temperature application, where creep is not important. On the other hand, for high temperature applications, fine-grained materials are to be avoided, as grain boundary sliding can add to creep deformation. There are no grain boundaries in a single crystal and, therefore, grain boundary sliding is not a problem here. Single crystal titanium turbine blades have been tried out, even though the cost is an inhibiting factor here. A less expensive way of minimizing grain boundary sliding is to orient the grains in such a way that the boundary sliding does not result in cavities at grain junctions. These cavities, if present, can coalesce together and cause creep fracture.

SUMMARY

1. The common tensile test on a ductile crystalline material gives data on yield strength, ultimate tensile strength, per cent elongation and per cent reduction in area.
2. Slip is the common mode of plastic deformation at ambient and elevated temperatures. Slip occurs along well defined crystallographic planes and directions.

3. The observed shear strengths of crystals are some three or four orders of magnitude smaller than the theoretical strengths. The motion of dislocations at low stresses accounts for this discrepancy. Only in perfect crystals such as whiskers, the theoretical strength is attainable.
4. The stress to move a dislocation in a crystal is determined by the bond strength and the bond directionality. Covalent crystals, requiring a large stress to move a dislocation, are not ductile. Typical metals need small stresses to move a dislocation and are ductile.
5. Crystals that are inherently hard show a strong dependence of yield stress on temperature.
6. There are sources for the generation of dislocations within crystals. These sources account for the increase in dislocation density with increasing amount of plastic deformation.
7. Work hardening is due to the formation of sessile (immobile) dislocations by the interaction of dislocations moving on non-parallel slip planes. The plastic flow strength increases during work hardening as the square root of the dislocation density.
8. Grain boundaries provide effective obstacles to dislocation motion. The yield stress of a polycrystalline material increases as the reciprocal of the square root of the grain diameter.
9. Dislocations interact with the stress fields of solute atoms, leading to solid solution strengthening. The hardness of martensite is an extreme example, where all the carbon dissolved in austenite is retained in martensite by quenching, resulting in substantial strengthening due to supersaturation.
10. Very closely spaced precipitate particles obstruct dislocation motion. The yield stress varies as the reciprocal of the particle spacing.
11. Creep is the thermally activated deformation that occurs as a function of time at temperatures above $0.4T_m$. Dispersion hardening is an effective method of improving creep resistance.

PROBLEMS

- 11.1** The yield stress, the strength coefficient K , and the work hardening exponent n for a quenched and tempered steel and for annealed copper are listed below.

Material	σ_y, M^{-2}	$K, \text{M m}^{-2}$	n
Steel	520	1270	0.15
Copper	55	317	0.54

Calculate the difference between the flow stress at 0.1 plastic strain and the yield stress for the two materials.

Answer: Steel: 379 and copper: 36 MN m^{-2} .

- 11.2** Find all the slip systems that have the following slip planes: (i) (111) in an FCC crystal and (ii) $(1\bar{1}0)$ in a BCC crystal.
- 11.3** Show that the resolved shear stress reaches a maximum value, when $\phi_1 = \phi_2 = 45^\circ$.
- 11.4** Sometimes, slip occurs on planes other than the close packed planes in HCP crystals. If the c/a ratio of Zn is 1.86 and that of Co is 1.62, which one is likely to slip more often on close packed planes and why?
- 11.5** A tensile stress of 10 MN m^{-2} is applied along the [112] direction of an iron (BCC) crystal. What is the shear stress in the [010] direction lying on the (001) plane?
Answer: 3.33 MN m^{-2} .
- 11.6** A zinc crystal (HCP) is oriented with normal to the basal plane making an angle of 60° with the tensile axis and the three slip directions x_1 , x_2 and x_3 lying on its plane making angles of 38° , 45° and 84° , respectively with the tensile axis. If the plastic deformation is first observed at a stress of 2.3 MN m^{-2} , find which of the three slip directions has initiated slip and at what value of the resolved shear stress?
Answer: x_1 at 0.9 MN m^{-2} .
- 11.7** An FCC crystal has a CRSS of 0.7 MN m^{-2} . What tensile stress must be applied along the [100] direction of the crystal to initiate plastic deformation?
Answer: 1.7 MN m^{-2} .
- 11.8** Estimate the width of dislocations in copper. The shear stress to initiate plastic deformation is $\mu/10^5$.
Answer: 4.7 \AA .
- 11.9** The activation volume for the intersection of a moving dislocation with other dislocations in the matrix can be as high as $5000 b^3$. Show that the dislocation intersection is essentially an athermal process at room temperature.
- 11.10** A cube of copper of $10 \text{ mm} \times 10 \text{ mm} \times 10 \text{ mm}$ is sheared at a rate of 10 mm per min . Estimate in order of magnitude the minimum number of dislocations in motion in this cube at a given instant. Assume a dislocation velocity of 1 km s^{-1} .
Answer: 10
- 11.11** Make a series of sketches to show how a Frank–Read source operates. How many loops are created from a F–R source in a single crystal of copper, when the corresponding slip displacement is 1 mm along the slip plane?
Answer: 4×10^6 .

- 11.12** The lengths of Frank–Read sources in a copper crystal vary from 10^{-8} m to 10^{-6} m. Estimate the shear stress at which plastic deformation is initiated.

Answer: 11 MN m^{-2} .

- 11.13** Compute the stored energy of copper cold worked from the annealed state to a dislocation density of 10^{14} m^{-2} .

Answer: 0.14 MJ m^{-3} .

- 11.14** In a cold-worked copper crystal, having a dislocation density of $3 \times 10^{14} \text{ m}^{-2}$, the shear stress to move a dislocation is found to be 100 MN m^{-2} . This value reduces to 1 MN m^{-2} on annealing. Find the shear stress to move a dislocation in a copper crystal cold worked to a dislocation density of $2 \times 10^{12} \text{ m}^{-2}$.

Answer: 9.08 MN m^{-2} .

- 11.15** Show that the average grain diameter decreases by half, for every increase of 2 in the ASTM number for grain size.

- 11.16** Estimate the yield stress of a polycrystalline Fe-3% Si alloy, when the grain size is ASTM 1, 4, and 8, respectively. Assume $\sigma_i = 80 \text{ MN m}^{-2}$ and $k = 0.63 \text{ MN m}^{-3/2}$.

Answer: 120, 147 and 214 MN m^{-2} .

- 11.17** What difficulty would you expect in increasing the concentration of the solute tin in bronze to 40% as in brass, for increasing the solid solution strengthening?

- 11.18** In a duralumin alloy, the precipitates are $1 \mu\text{m}$ diameter and the average distance between centres of neighbouring particles is $10 \mu\text{m}$. Estimate the yield stress of the alloy. Is this alloy aged to the optimum extent? Explain.

Answer: 0.79 MN m^{-2} .

- 11.19** Supersonic aircraft with a speed greater than Mach 2 attain a skin temperature of over 250°C . Explain why a duralumin alloy cannot be used for such an application.

- 11.20** Compare the methods available for increasing the low temperature strength and the creep strength of a crystalline solid.

- 11.21** In an FCC crystal, the tensile axis is along the $[12\bar{1}]$ direction. Determine the slip system that is most favourably oriented for slip to start at the lowest applied tensile stress.

$$[\text{Hint: } \cos \phi = \frac{h_1 h_2 + k_1 k_2 + l_1 l_2}{\sqrt{(h_1^2 + k_1^2 + l_1^2)(h_2^2 + k_2^2 + l_2^2)}}]$$

where ϕ is the angle between two planes or two directions.]

Answer: $(111) [01\bar{1}]$ or $(\bar{1}1\bar{1}) [110]$

- 11.22** Compute the mean grain diameters in mm, corresponding to ASTM grain size numbers (i) 3, (ii) 0, and (iii) -2.5.
Answer: 0.125 mm, 0.35 mm, and 0.84 mm.
- 11.23** A grain size of ASTM 7 in a mild steel is refined to ASTM 14 by the addition of microalloying elements. Estimate the increase in the yield strength of the steel.
Answer: 300 MN m⁻².
- 11.24** Two samples of an Fe-3%Si alloy have grain sizes of ASTM 1 and 8 respectively. They have yield strengths of 118 and 207 MN m⁻². Calculate the yield strength of a single crystal of this alloy.
Answer: 81 MN m⁻².
- 11.25** Estimate the yield strength of a cold-worked microalloyed steel from the following data:
 Peierls–Nabarro (tensile) stress for Fe : 35 MN m⁻²
 Solid solution strengthening by Mn and Si : 135 MN m⁻²
 Dislocation density due to cold working : 10¹⁴ m⁻²
 ASTM grain size number : 13
 Assume that the constant A in the work hardening equation is 10 N m⁻¹.
Answer: 630 MN m⁻².
- 11.26** Describe briefly the main strengthening methods against plastic yield at low temperatures.
- 11.27** Reexamine critically the above strengthening methods against plastic yield for temperatures above $0.4T_m$.
- 11.28** Discuss the role of (i) grain boundaries and (ii) precipitate particles, in strengthening crystalline materials against yield at low as well as high temperatures of deformation.

MULTIPLE CHOICE QUESTIONS

- As compared to engineering stress-strain curve, the true stress-strain curve is
 - above and to the left
 - below and to the right
 - crosses the engineering curve
 - parallel to the engineering curve
- In a tensile test, the engineering stress corresponding to the maximum load is called
 - yield strength
 - tensile strength
 - UTS
 - upper yield stress

3. In a tensile test, necking starts at
 - A. lower yield stress
 - B. upper yield stress
 - C. ultimate tensile stress
 - D. just before fracture
4. Superplastic materials have an index of strain-rate sensitivity m in the range of
 - A. 0
 - B. 0.1–0.2
 - C. 0.4–0.9
 - D. 1.5–2.0
5. The critical resolved shear stress (CRSS) for a polycrystalline copper is
 - A. $\mu/6$
 - B. $\mu/88,000$
 - C. $\mu/110,000$
 - D. not defined
6. The resolved shear stress for plastic deformation to start in an iron crystal is (μ = shear modulus)
 - A. μ
 - B. $\mu/6$
 - C. $\mu/30$
 - D. $\mu/4700$
7. Crystals like diamond and silicon are brittle, because
 - A. they contain no dislocations
 - B. they are noncrystalline
 - C. the stress to move a dislocation is high in them
 - D. they contain very few dislocations
8. Copper is ductile, because
 - A. it is a perfect crystal
 - B. it contains a very high density of dislocations
 - C. it has glassy structure
 - D. the stress to move a dislocation in it is low
9. For copper, the yield stress σ_y and the brittle fracture stress σ_f are related as
 - A. $\sigma_y > \sigma_f$
 - B. $\sigma_f > \sigma_y$
 - C. $\sigma_f < \sigma_y$
 - D. $\sigma_y = \sigma_f$
10. The length of the Frank–Read source operating after work hardening to double the initial yield stress is approximately
 - A. half
 - B. same as before
 - C. double
 - D. 4 times
11. The yield strength of an annealed copper crystal is 1 MPa. It increases to 9 MPa on cold working to a dislocation density of $2 \times 10^{12} \text{ m}^{-2}$. The value of the constant A in the equation describing the work-hardening effect is
 - A. 5.66 N m^{-1}
 - B. 5.66 N m^{-2}
 - C. 5.72 N m^{-1}
 - D. $3 \times 10^{-13} \text{ N m}^{-2}$
12. The yield stress of 1 MN m^{-2} of an annealed copper crystal increases to 100 MN m^{-2} on cold working to a dislocation density of 10^{14} m^{-2} . If cold working had been done to a dislocation density of 10^{12} m^{-2} , the yield stress would be
 - A. 109 MN m^{-2}
 - B. 67 MN m^{-2}
 - C. 10.9 MN m^{-2}
 - D. no change

13. The yield stress of a copper crystal is 100 and 10 MN m⁻², when cold worked to a dislocation density of 3×10^{14} and 2×10^{12} m⁻², respectively. The yield stress of the softest crystal is
A. 0 MN m⁻² B. 2 MN m⁻² C. 10 MN m⁻² D. 3.33 MN m⁻²
14. If ASTM grain size number 7 corresponds to grain diameter of 0.03 mm, ASTM 4 corresponds to
A. 0.24 mm B. 0.085 mm C. 0.011 mm D. 0.004 mm
15. A grain size measured wrongly at a magnification of $300 \times$ gave an ASTM number of 2. The correct grain size number should be
A. -1 B. 5.2 C. 6.0 D. 5.0
16. The mean grain diameter corresponding to ASTM number = -0.5 is
A. 5 mm B. 0.05 mm C. 0.43 mm D. 0.30 mm
17. The ASTM grain size number of a material which has 28 grains per sq.in. at a magnification $75 \times$ is about
A. 14 B. 7 C. 5 D. 4
18. A mild steel ($k = 0.7$ MN m^{-3/2}) of 0.03 mm grain diameter is grain refined to 0.003 mm. The increase in yield strength (MN m⁻²) will be about
A. 275 B. -275 C. -0.00275 D. 8.1
19. The yield stresses of an Fe-4%Si alloy at grain sizes of 0.015 and 0.045 mm are 214 and 147 MN m⁻². The yield stress of a single crystal is
A. 55.5 MN m⁻² B. 120 MN m⁻² C. 80 MN m⁻² D. 360 MN m⁻²
20. The yield stresses of both Au and Ag single crystals are 0.7 MPa. The yield stress of a 50 : 50 solid solution is
A. 0.7 MPa B. > 0.7 MPa C. < 0.7 MPa D. 0
21. The radii of copper, zinc and tin are 1.28, 1.31 and 1.51 Å respectively. As compared to zinc, tin in copper will produce a strengthening effect that is
A. large B. small C. very small D. same
22. When the inter-precipitate spacing increases from 200 to 2000 Å, the ratio of the new yield strength to the initial strength is
A. 10 B. 1 C. 0.1 D. none of these
23. The methods to increase the yield strength of a crystalline material are
A. grain refinement B. annealing
C. solute additions D. precipitation hardening
24. The most desirable method of increasing the yield strength of mild steel is
A. grain refinement B. cold working
C. solute additions D. precipitation hardening

25. If the melting point of polyethylene is 140°C , it will not creep at
 A. room temperature B. 0°C
 C. -60°C D. -196°C
26. With melting points given in brackets, tick those materials which will creep significantly at 180°C
 A. Pb (327°C) B. Cu (1084°C) C. Al (660°C) D. W (3410°C)
27. TD (thoria dispersed) nickel has adequate creep resistance up to $0.7 T_m$ because
 A. thoria has a high melting point
 B. thoria does not dissolve in nickel
 C. nickel has a high melting point
 D. nickel gets work hardened during service

Answers

- | | | | | |
|----------|---------|-------------|-------|-------|
| 1. A | 2. B, C | 3. C | 4. C | 5. D |
| 6. D | 7. C | 8. D | 9. B | 10. A |
| 11. A | 12. C | 13. B | 14. B | 15. B |
| 16. C | 17. C | 18. A | 19. A | 20. B |
| 21. A | 22. C | 23. A, C, D | 24. A | 25. D |
| 26. A, C | 27. B | | | |

Sources for Experimental Data

- ASM International, *Metals Handbook*, 10th ed., Vol. 1, *Irons and Steels and High Performance Alloys: Specialty Steels and Heat Resistant Alloys*, pp. 755–1003, Materials Park, Ohio (1990).
- A. Kelly and R.B. Nicholson, *Strengthening Methods in Crystals*, Elsevier, Amsterdam (1971).

Suggestions for Further Reading

- R.W.K. Honeycombe, *Plastic Deformation of Metals*, Edward Arnold, London (1984).
- W.D. Kingery, H.K. Bowen and D.R. Uhlmann, *Introduction to Ceramics*, John Wiley, New York (1976), Chap. 14.

CHAPTER

12

Fracture

Fracture refers to the failure of a material under load by breaking into two or more pieces. Fracture can occur under all service conditions. Materials subjected to alternating or cyclic loading (as in machines) fail due to fatigue. The fracture under such circumstances is called fatigue fracture. Materials used at high temperatures can fail due to creep fracture. The characteristics of the various fractures have been studied extensively. Preventing failure during service is one of the most important problems facing the engineer.

In this chapter, we briefly discuss the nature of ductile and brittle fracture and the fracture toughness of materials. We describe the methods of protecting materials against fracture and fatigue failure at the end of the chapter.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Fracture stress σ_f	meganewton per square metre	MN m ⁻²	kgf/mm ² , dyne/cm ² , psi, ksi
Crack length c	micrometre	μm	inch, cm
Surface energy γ	joule per square metre	J m ⁻²	erg/cm ²
Young's modulus Y	giganewton per metre ²	GN m ⁻²	kgf/mm ² , dyne/cm ²
Fracture toughness G_c	joule per square metre	J m ⁻²	lb-in/in ²
Fracture toughness K_{Ic}	meganewton per (metre) ^{3/2}	MN m ^{-3/2}	ksi $\sqrt{\text{in}}$

12.1 Ductile Fracture

Ductile fracture is the rupture of a material after a considerable amount of plastic deformation. Materials begin to neck beyond the ultimate tensile strength, which is at the maximum point in the load-elongation curve, see Fig. 11.1a. The neck refers to the reduced cross-section of the specimen near the middle. The

true stress in this region is increasing, in spite of the fall in the load and the engineering stress. *Fully ductile materials* will continue to neck down to an infinitesimally thin edge or a point and thus fail, as the cross-section at the neck becomes so small that it cannot bear the load any longer. The more common type of ductile fracture occurs when the reduced cross-section has still an appreciable area. Here, *the cracks are found to nucleate at brittle particles*: either the natural kind found in multiphase materials, e.g., cementite in steel, or foreign inclusions, e.g., oxide inclusions in copper. When a brittle particle is present, it is difficult to maintain compatibility in the neck region between the continuously deforming matrix and the nondeforming particle. This results in the formation of very tiny voids near the matrix-particle interface. If fracture initiates at pores in the neck region, then the voids are already present. The voids grow with increasing deformation and ultimately reach sizes of the order of a mm. At this stage, the material may tear apart, see Fig. 12.1. The effective

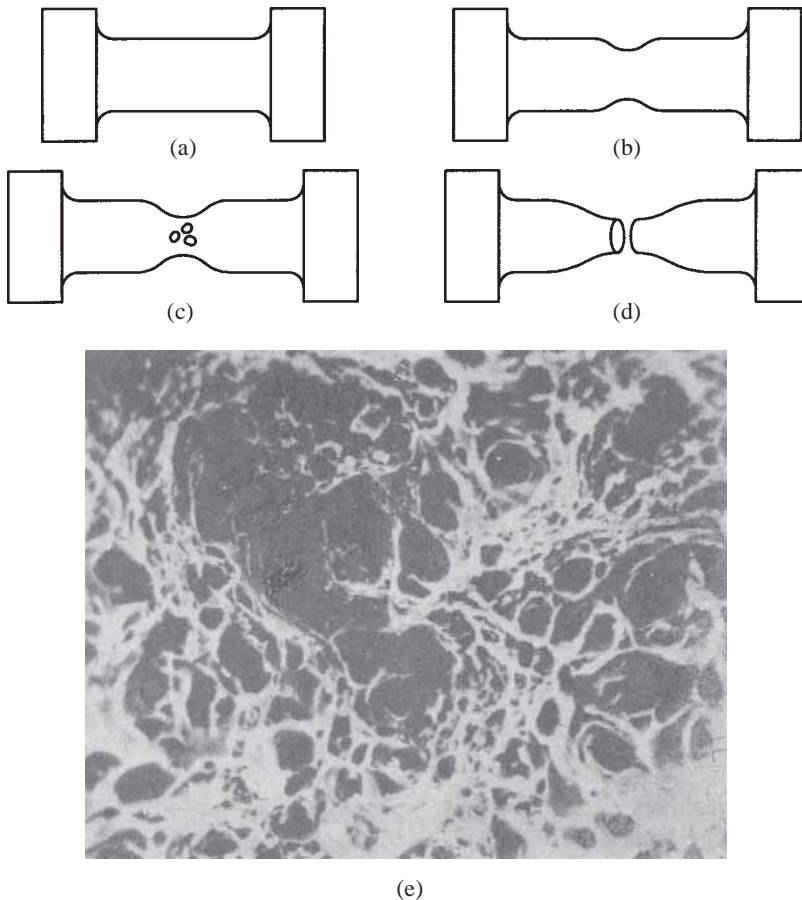


Fig. 12.1 (a–d) Successive stages in the ductile fracture of a tensile test specimen. (e) Scanning electron micrograph of the fractured surface of an aluminium alloy, showing dimples suggestive of void growth. [(e) A.S. Argon (Ed.), *Physics of Strength and Plasticity*, by permission from the MIT Press, Cambridge, Mass.]

cross-section of the neck is small due to the presence of these discontinuities, even though the apparent cross-section may be appreciable. This model of ductile fracture is supported by the finding that materials of ultra high purity which do not have inclusions or pores, do rupture in a fully ductile manner, that is, after thinning down to a point or an edge.

12.2 Brittle Fracture

Brittle fracture is the failure of a material *without apparent plastic deformation*. If the broken pieces after a brittle fracture are fitted together, the original shape and dimensions of the specimen are restored. In an ideal material, fracture can be visualized as the pulling apart and breaking of the interatomic bonds across two neighbouring atomic planes. A simple calculation similar to that done for the theoretical shear strength of a perfect crystal (Sec. 11.3) shows that the tensile stress required to break the interatomic bonds across two adjacent atomic planes is of the order of $Y/6$, where Y is the Young's modulus of the material. Brittle materials, however, break at a much lower stress, of the order of $Y/1000$. This difference of two orders of magnitude between the calculated and the observed strengths is similar to the discrepancy between the strengths of perfect and real crystals, see Sec. 11.3. Just as the presence of dislocations in real crystals explains the discrepancy, here the existence of tiny cracks in brittle materials is the reason for their poor tensile strength.

Griffith has postulated a criterion for the propagation of a pre-existing crack in a brittle material. Consider a lens shaped crack of length $2c$ shown in Fig. 12.2. The material is of unit thickness and the crack runs from the front to

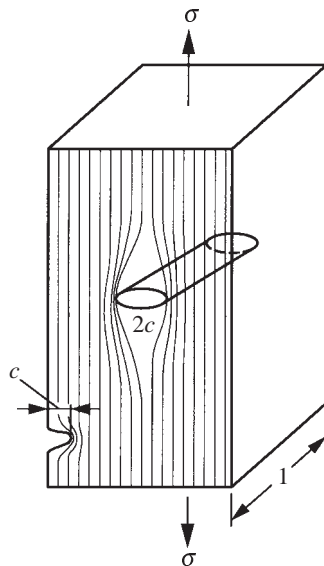


Fig. 12.2 Illustration of Griffith's criterion. The flow lines depict elastic lines of force.

the back face. When a longitudinal tensile stress σ is applied, the crack tends to increase its length in the transverse direction. If the crack spreads, the surface area of the crack increases, while the elastic strain energy stored in the material decreases; this is because the elastic strains cannot be continuous across the cracked region. If γ is the surface energy per unit area of the material, then the surface energy of the crack is $4\gamma c$. Note that two surfaces are created, as the crack spreads. The elastic energy density (elastic energy per unit volume of the material) is $(1/2)\sigma\epsilon$, where ϵ is the elastic strain. Substituting for ϵ in terms of the Young's modulus Y , the elastic energy density becomes $(1/2)\sigma^2/Y$. As a first approximation, we can take that no elastic energy is stored in a cylindrical volume around the crack, where there are no lines of force in Fig. 12.2. Thus the elastic energy released, when the crack is introduced, is equal to $\pi\sigma^2c^2/2Y$. A more detailed calculation shows that this strain-free volume is, in fact, larger. It is better approximated to twice the volume of the cylindrical region. Correspondingly, the elastic energy release is also twice the above value.

If we write an energy equation for the crack formation, the energy change ΔU as the crack forms is given by

$$\Delta U = 4\gamma c - \frac{\pi\sigma^2c^2}{Y} \quad (12.1)$$

The negative sign of the second term indicates that the elastic energy stored in the material is released, as the crack forms. This is opposite in sense to the energy consumed in creating the two surfaces of the crack. In Fig. 12.3, ΔU is plotted as a function of c , showing that, for a given value of σ , ΔU passes through a maximum at a critical value of c equal to c^* . The larger is the value of

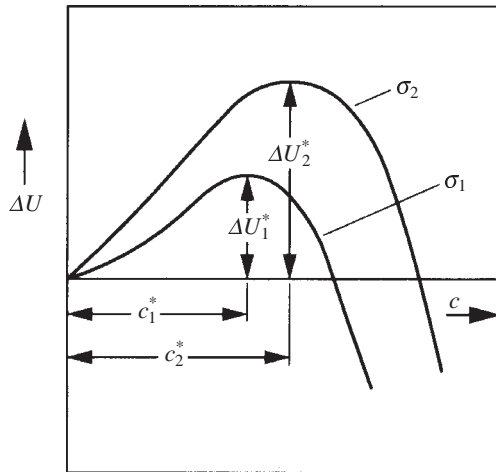


Fig. 12.3 The energy ΔU as a function of the crack half-length c for two tensile stresses σ_1 and σ_2 ($\sigma_1 > \sigma_2$).

σ , the smaller is c^* . The critical value can be found by setting $d\Delta U/dc = 0$. As the applied tensile stress is the external variable for a given material with a pre-existing crack of length $2c$, it is appropriate to express the critical condition as a critical fracture stress σ_f :

$$\sigma_f = \sqrt{\frac{2\gamma Y}{\pi c}} \quad (12.2)$$

When the critical stress is applied to a brittle material, the pre-existing crack propagates spontaneously with a decrease in energy, culminating in fracture. Thus, Eq. (12.2) gives Griffith's energy balance criterion for crack propagation.

Example 12.1 A sample of glass has a crack of half length $2 \mu\text{m}$. The Young's modulus of the glass is 70 GN m^{-2} and the specific surface energy is 1 J m^{-2} . Estimate its fracture strength and compare it with its Young's modulus.

Solution Substituting the given values of the crack half-length, the Young's modulus and the surface energy in Eq. (12.2), we obtain

$$\begin{aligned} \sigma_f &= \sqrt{\frac{2 \times 1 \times 70 \times 10^9}{3.14 \times 2 \times 10^{-6}}} \\ &= 150 \text{ MN m}^{-2} \end{aligned}$$

The fracture strength of 150 MN m^{-2} is 1/500th of the Young's modulus, which is 70 GN m^{-2} . Thus the Griffith's criterion bridges the gap between the observed and the ideal strengths of brittle materials.

The Griffith equation (12.2) is equally valid for a surface crack of length c . So, a surface crack of depth c (see Fig. 12.2) is as effective as an internal crack of length $2c$. Hence, in a given material, surface cracks are more effective than internal cracks. A brittle material held with grips that do not extend up to the surface usually shows a higher strength. Here, the surface is free of stresses. The surface cracks are consequently ineffective.

All materials have a statistical distribution of cracks in them. The longest crack that is most favourably oriented with respect to the stress axis is the most effective one. It, therefore, propagates *first* to cause fracture, as the applied stress is increased. If a second test is carried out on one of the broken pieces, the strength is usually higher as a direct consequence of the elimination of the most effective crack in the first test.

A crack may not propagate even after the Griffith condition is satisfied, if there is not sufficient stress concentration at the crack tip. Consider the schematic crack tip depicted in Fig. 12.4 on the atomic scale. The propagation of the crack obviously requires the breaking of the interatomic bonds at the tip of

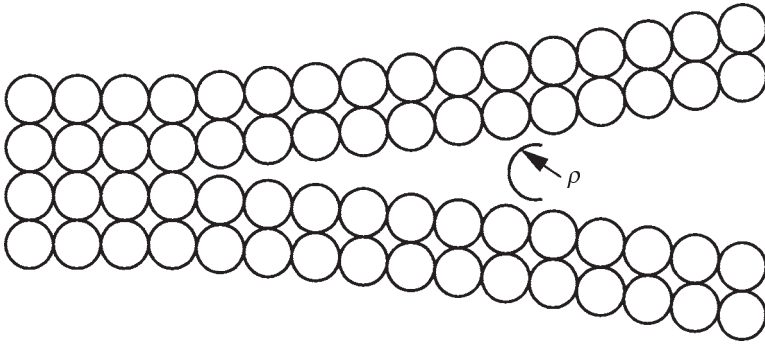


Fig. 12.4 An atomically sharp crack has a radius of curvature ρ at its tip equal to about $1\text{--}2 \text{ \AA}$.

the crack. This bond rupture would be feasible only if there is a stress concentration at the tip equal to the theoretical stress required to pull apart the atomic bonds. The stress concentration σ_{\max} at the tip of a crack, when a stress σ is applied, is given by

$$\sigma_{\max} = 2\sigma\sqrt{c/\rho} \quad (12.3)$$

where ρ is the radius of curvature at the tip of the crack. Note that the lines of force, shown in Fig. 12.2, are crowded at the tip, indicating a stress concentration there. For the atomically sharp crack shown in Fig. 12.4, ρ is about 1 \AA (10^{-10} m). With $c \sim 1 \text{ }\mu\text{m}$, this yields $\sigma_{\max} = 200\sigma$. This concentrated stress is of the same order as the ideal strength ($Y/6$) and, therefore, is sufficient to cause rupture of the atomic bonds at the tip.

In brittle materials such as silicate glass, plastic deformation is nonexistent. The stress concentration at the tip of the crack is not relaxed due to plastic deformation and consequently the crack tip remains sharp. The term γ in the Griffith equation is truly representative of the surface energy of the crack faces, as no other additional work such as plastic work is done during the propagation of the crack. In such cases the Griffith's criterion for crack propagation is fully valid.

In many crystalline materials, fracture occurs in an apparently brittle manner where the broken pieces can be fitted together to restore the original shape and dimension of the specimen. However, on a microscopic scale, some plastic deformation always occurs in a few grains in the neighbourhood of the crack tip, where the stress is concentrated. This deformation has a two-fold effect. It increases the work that is necessary to propagate a crack, because the work now includes not only the surface energy of the crack faces but also the energy consumed in plastically deforming the matrix adjacent to the crack tip. The other effect is to relax the stress concentration at the tip, making the crack blunt. Both these effects tend to make the crack propagation more difficult, thereby increasing the fracture resistance of the material.

Attempts have been made to take into account the plastic work, by using a modified form of the surface energy term called γ_p in the Griffith equation. In a number of cases, the plastic work term is much larger than the surface energy term, requiring much longer cracks or much higher stresses to reach the critical condition in a ductile material.

Example 12.2 The half length of cracks in a steel is 2 μm . Taking $Y = 200 \text{ GN m}^{-2}$, estimate the brittle fracture strength at low temperatures, if the true surface energy is 1.5 J m^{-2} . The actual fracture strength is found to be 1200 MN m^{-2} . Explain the difference, if any, between this and your result.

Solution From Griffith's equation,

$$\begin{aligned}\sigma_y &= \sqrt{\frac{2\gamma Y}{\pi c}} = \sqrt{\frac{2 \times 1.5 \times 200 \times 10^9}{3.14 \times 2 \times 10^{-6}}} \\ &= 310 \text{ MN m}^{-2}\end{aligned}$$

The observed strength is four times larger than the calculated value. As iron plastically deforms, the higher observed strength can be attributed to the plastic work done by the crack as it propagates. Resubstituting the observed strength in the Griffith equation, the effective surface energy that includes the plastic work is 16 times larger than the true surface energy.

12.3 Fracture Toughness

Defects, cracks or flaws are inevitably present in all engineering materials. They may be introduced during solidification, fabrication or heat treatment stages of the material. The fracture resisting capability of a machine component or an engineering structure therefore must be evaluated in the presence of cracks. The fracture resistance of a material in the presence of cracks or discontinuities is known as its *fracture toughness*.

From the Griffith type of approach, the fracture toughness is defined by the critical value of a parameter G_c . G_c gives the value of *the strain energy released per unit area of the crack surface* when unstable crack extension (leading to fracture) takes place. For an elastic crack of length $2c$, the critical strain energy release rate G_c is equal to $\sigma_f^2 \pi c / Y$. A comparison of this with Eq. (12.2) shows that

$$G_c = 2\gamma \quad (12.4)$$

In an ideally brittle material such as a silicate glass, G_c can be equated to 2γ . In materials, where plastic deformation occurs during crack initiation, G_c can be much larger than 2γ . Since the plastic work term is dominant as compared to the true surface energy term, G_c for such materials may be written as

$$G_c = 2\gamma_p \quad (12.5)$$

For example, G_c for an aluminium alloy is in the range 20–100 kJ m⁻², but the true surface energy γ is less than 2 J m⁻². Polystyrene, even though relatively brittle as compared to polyethylene, has G_c of about 3 kJ m⁻² and γ of only about 1 J m⁻².

Another parameter, which is more commonly used to describe the fracture toughness of a material, is known as the *critical stress intensity factor* K_{Ic} (or K_c). For a sharp crack in an infinitely wide plate, when the applied tensile stress is perpendicular to the crack faces, the critical stress intensity factor is given by

$$K_{Ic} = \sigma_f \sqrt{\pi C} \quad (12.6)$$

Fracture initiates in a material as soon as K_{Ic} is reached, either through increasing stress or increasing c or both. When the width of the material is finite or the loading geometry is not as stated above, a geometrical factor α is added to the right side of Eq. (12.6). K_{Ic} is a material property, just as yield strength, UTS, etc., and can be determined from fracture tests. K_{Ic} for mild steel has been determined to be about 25 MN m^{-3/2}, which is much larger than that for a fully elastic fracture.

In fracture safe design, if the size of the most damaging defect is known, the design stress can be computed from K_{Ic} . Alternatively, if the design stress is given, the critical crack size calculated from K_{Ic} must be sufficiently larger than the smallest size of the crack detectable by available inspection techniques.

12.4 The Ductile-Brittle Transition

Common BCC metals become brittle at low temperatures or at extremely high rates of strain. Many FCC metals, on the other hand, remain ductile even at very low temperatures. When the slip systems on the basal plane are the only ones operating, polycrystalline HCP metals are brittle, as there are not enough slip systems to maintain the grain boundary integrity. For the same reason, polycrystalline ionic crystals are also brittle.

The conditions under which a material behaves in a brittle fashion depend on several factors. BCC metals generally require a high stress to move dislocations and this stress increases rapidly with decreasing temperature. The stress required to propagate a crack, on the other hand, is not a strong function of temperature. So, at some temperature called the *ductile-brittle transition temperature*, the stress to propagate a crack, σ_f , is equal to the stress to move dislocations, σ_y . At temperatures higher than the transition temperature, $\sigma_y < \sigma_f$ and the material first yields plastically. At temperatures lower than the transition temperature, the material is brittle. Here, the actual brittle fracture stress may be controlled by the yield stress, as some microscopic yielding may be necessary to nucleate a crack. So, at all temperatures below the transition temperature, $\sigma_f = \sigma_y$. As soon as the applied stress reaches σ_y , the crack is nucleated at the intersection of slip planes and propagates rapidly.

Figure 12.5 shows the impact test results of steels of different carbon content as a function of temperature. The energy to break a specimen under

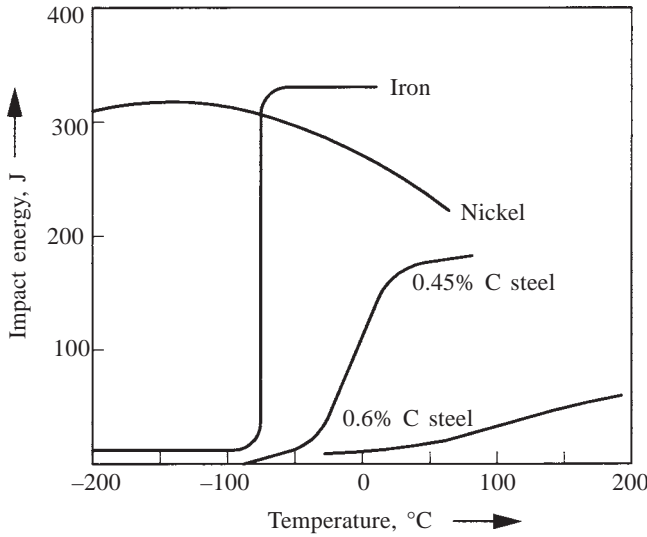


Fig. 12.5 The variation of impact energy as a function of temperature for different materials.

impact loading decreases abruptly at a certain temperature, which corresponds to the transition from ductile to brittle behaviour.

The yield stress σ_y is a function of grain size, recall the Hall–Petch equation (11.22). The yield strength increases with decreasing grain size. So, the stress to nucleate a crack by slip intersection increases with decreasing grain size. The stress to propagate a crack σ_f is also a function of grain size, as the length $2c$ of the nucleating crack can be equated to the grain diameter. The stress to propagate a crack therefore increases with decreasing grain diameter. Thus, fine-grained materials have a lower transition temperature as compared to coarse-grained materials. The transition temperature is a function of several other variables. Sharp notches in the specimen provide stress concentration centres and hence increase the transition temperature. A high strain rate increases the stress needed to move a dislocation and, therefore, increases the transition temperature.

Example 12.3 The temperature and strain rate dependence of yield stress in MN m^{-2} for molybdenum is given by

$$\sigma_y = 20.6 + 173 \, 600/T + 61.3 \log_{10} (d\varepsilon/dt)$$

where T is the temperature in kelvin and $(d\varepsilon/dt)$ is the strain rate in s^{-1} . Sharp cracks of half length $2 \, \mu\text{m}$ are present in the metal. Estimate the temperature at which the ductile to brittle transition occurs at a strain rate of (i) $10^{-2} \, \text{s}^{-1}$ and (ii) $10^{-5} \, \text{s}^{-1}$. $Y = 350 \, \text{GN m}^{-2}$ and specific surface energy is $2 \, \text{J m}^{-2}$.

Solution At the transition temperature, the yield stress should be equal to the fracture stress.

From Eq. (12.2), the fracture stress is

$$\sigma_f = \sqrt{\frac{2 \times 2 \times 350 \times 10^9}{3.14 \times 2 \times 10^{-6}}} = 470 \text{ MN m}^{-2}$$

(i) At a strain rate of 10^{-2} s^{-1} , the yield stress is

(ii) $\sigma_y = 20.6 + 173 \, 600/T + 61.3 \times (-2)$

Setting $\sigma_y = \sigma_f$, the transition temperature can be calculated as

$$T = 173 \, 600/(470 - 20.6 + 122.6) = 300 \text{ K}$$

(iii) Setting again $\sigma_y = \sigma_f$, for a strain rate of 10^{-5} s^{-1} , we have

$$T = 173 \, 600/(470 - 20.6 + 306.5) = 230 \text{ K}$$

Polymeric materials are brittle below the glass transition temperature T_g . Molecular motion is frozen below T_g , so that viscous flow or high elasticity becomes restricted. The brittle behaviour of polymers under such conditions can be described following the Griffith approach.

The behaviour of ionic crystals is strongly dependent on their surface condition. Crystals free of surface defects exhibit limited plastic deformation caused by dislocation motion. But the presence of surface flaws makes the crystals prone to brittle failure. Under these conditions, even single crystals can be brittle.

12.5 Fracture Mechanism Maps

Attempts have been made mainly by M.F. Ashby and his co-workers to classify the large amount of experimental data on fracture characteristics of different groups of materials. The general correlations identified by them are plotted in what are called fracture mechanism maps. On these maps (Fig. 12.6), the mechanisms are shown for different groups of materials as a function of T/T_m and of the normalised tensile stress (tensile stress divided by the Young's modulus). Here, the term transgranular fracture refers to the propagation of fracture across grains. Intergranular (IG) fracture is fracture along the grain boundaries.

In FCC metals (Fig. 12.6a), the brittle fracture is absent at all temperatures. The ductile fracture is dominant at lower temperatures and creep failure at higher temperatures. The creep fracture can be transgranular or intergranular. Rupture refers to a fully-ductile failure at over 90% of reduction in area.

In refractory BCC metals such as W, Mo, Ta, Nb and Cr (Fig. 12.6b), at low temperatures, brittle intergranular fracture (BIF) is dominant. In Fig. 12.6, BIF I refers to brittle fracture by the propagation of preexisting cracks. BIF II refers to brittle fracture caused by cracks nucleated by the initiation of slip or twinning. BIF III refers to fracture by crack propagation after a significant amount of plastic strain (1–10%) has occurred. As the temperature increases, refractory BCC metals exhibit ductile fracture, creep fracture and rupture at successively higher temperatures.

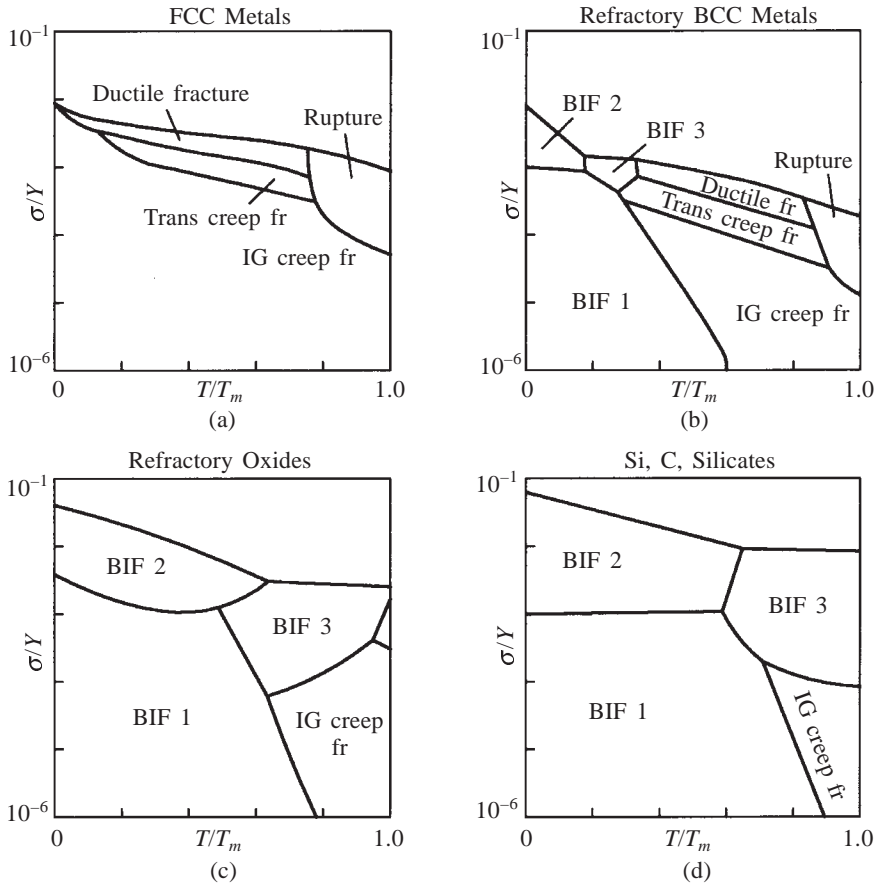


Fig. 12.6 Fracture mechanism maps.

In refractory oxides such as MgO , Al_2O_3 and Si_3N_4 (Fig. 12.6c), the brittle intergranular fracture (BIF) dominates the temperature range, the creep fracture appearing only at very high temperatures. Again, in covalently-bonded materials, such as C, Si and the silicates (Fig. 12.6d), the BIF region is dominant over almost the entire temperature range.

12.6 Methods of Protection against Fracture

We have already seen that a surface crack of length c is as effective as an internal crack of length $2c$. In addition, for a crack exposed to the atmosphere, the surface energy term γ can be effectively lowered by chemical adsorption of molecules on the crack faces. This in turn reduces the critical fracture stress.

Surface treatment is the most important method used to protect against brittle fracture. The fact that surface cracks are the ones which cause failure can be shown by several experiments. When the surface of a glass is etched with

hydrofluoric acid, the surface layers and the cracks in them are removed. Consequently, a freshly etched glass has a higher strength than the unetched glass. If this fact is to be exploited as a method of strengthening, then the etched surface should be protected against further mechanical abrasion that is likely to occur during service, as well as against the environmental effects such as corrosion. The surface of freshly drawn glass fibres is covered with a resin for protection. In addition, the drawing of glass into thin filaments or fibres has the effect of elongating any internal cracks in the direction of pulling. When a stress is applied parallel to the fibre length, the elongated cracks are ineffective, as there is no tensile stress perpendicular to the crack faces. Glass fibres of very high strengths have been produced and are used for strengthening a soft matrix in a composite material. During service, if some fibres break accidentally, there is no catastrophic failure, as the crack through the fibre gets blunted on reaching the soft matrix.

The surface cracks can be made ineffective also by other methods. If a *compressive stress* is introduced at the surface, the tensile stress required to cause the surface cracks to propagate is increased by a magnitude equal to the compressive stress. This method is adopted in the production of tempered glass used for shop windows and automobile windshields. The silicate glass is heated above its softening temperature and annealed long enough to remove all residual stresses. Then a blast of cold air is directed against the glass, so that the surface layers cool rapidly. They contract and become rigid. Glass being a poor thermal conductor, the heat conduction from the inside to the surface is slow. The interior is still above the softening temperature and is capable of flowing viscously to compensate for the contraction at the surface. There are no residual stresses introduced at this stage. Eventually, the inside is also cooled to room temperature. The associated contraction is now resisted by the rigid outer layers, so that compressive stresses are introduced in the outer layers and tensile stresses in the interior. This heat treatment called *tempering* results in better resistance to crack propagation. The compressive stresses can increase the fracture strength of the glass by 2–3 times. A faulty heat treatment, however, can introduce tensile stresses at the surface, making the glass more prone to fracture after heat treatment.

Compressive stresses can also be introduced in the surface layers by the *ion exchange method*, where smaller cations such as Na^+ of a sodium silicate glass are replaced at the surface by larger cations such as K^+ , by a chemical reaction at the surface. This chemical strengthening can introduce compressive stresses, which are of larger magnitude, compared to that obtained by the tempering method.

Grinding of brittle ceramic materials usually introduces cracks near the surface, which extend up to a depth of one grain diameter. Hence, fine grain sizes effectively decrease the size of the surface cracks and are desirable for improved strength of the ceramic. The very fine grain sizes (about $0.1\ \mu\text{m}$) obtained in crystallized glass (pyroceram) accounts for its excellent thermal and mechanical shock resistance.

In many metals, *intergranular failure*, that is, failure along the grain boundaries, occurs, if there is a continuous brittle phase present at the boundaries. In steels, sulphur can be present at the grain boundaries as a thin brittle iron sulphide film. This dangerous distribution can be prevented if there is enough manganese in the steel to form spherical manganese sulphide particles within the grains in place of the continuous film at the boundary.

Example 12.4 Iron tested at -196°C is brittle. It, however, becomes ductile at this temperature, if a thin layer of silver is diffused along the grain boundaries of iron. Explain this change in behaviour.

Solution Silver being an FCC metal has a low Peierls–Nabarro stress and hence is ductile at -196°C . The ductile layer of silver at the grain boundaries of iron effectively prevents crack propagation from one grain of iron into a neighbouring grain. With the cracks not propagating, the stress can be increased sufficiently to initiate plastic deformation in iron.

12.7 Fatigue Fracture

Rotating shafts, connecting rods, aircraft wings and leaf springs are some examples of structural and machine components that are subjected to millions of cycles of alternating stresses during service. The majority of failures of such components in service are due to fatigue. A fatigue failure can occur even below the yield stress of a material. For example, the yield strength of mild steel is 220 MN m^{-2} , but it will fail at a stress of 140 MN m^{-2} , if it is subjected to a very large number of stress reversals.

The fatigue behaviour of a material is understood from the results of a fatigue test, which are presented in the form of S – N curves. Samples of the material are subjected to alternating stresses of different levels. The number of cycles of stress reversals N required to cause fracture is plotted against the applied stress level S as shown in Fig. 12.7.

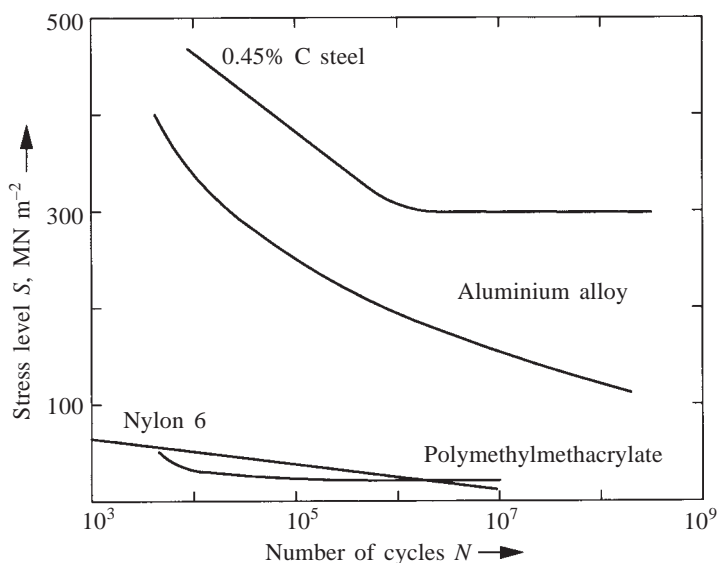


Fig. 12.7 The S – N curves for different materials.

Some materials such as mild steel show a clearly defined *fatigue limit*. If the applied stress is below the fatigue limit, the material will withstand any number of stress reversals. Other materials do not show a clearly defined limit. Here, the fatigue limit (or the endurance limit) is defined as the stress that would cause failure after a specified number of stress reversals.

Fatigue fracture occurs by crack propagation. The crack usually initiates at the surface of the specimen and propagates slowly at first into the interior. At some critical stage, the crack propagation becomes rapid culminating in fracture.

The fatigue life of a component can be improved by several methods. *Good design* plays an important role in increasing the fatigue life. Sharp corners should be avoided in design, so that regions of stress concentration are not present. *Polishing* the surface of the component to a good finish also removes some of the surface irregularities, which may initiate a crack. *Shot peening* of metals introduces compressive stresses at the surface and improves the fatigue strength. *Carburizing and nitriding* introduce strong surface layers and increase the resistance to crack initiation at the surface. On the other hand, decarburization produces a soft surface layer that lowers the fatigue resistance. A *fine grain size* improves the fatigue resistance.

SUMMARY

1. Prevention of fracture in materials during fabrication or service is an important problem for the engineer.
2. Ductile fracture is commonly caused by the formation and coalescence of pores or voids in the necked region.
3. Brittle fracture is usually caused by the propagation of pre-existing cracks in a material. Griffith's criterion gives the critical stress required to propagate a crack spontaneously.
4. In truly brittle materials, cracks are atomically sharp, so that there is sufficient stress concentration at their tips to break open the bonds during crack propagation.
5. In ductile materials, plastic deformation at the crack tip blunts the cracks as well as increases the energy of propagation of the crack.
6. Many materials exhibit a transition to brittle behaviour at low temperatures. Surface perfection, grain size, the stress required to move a dislocation and molecular mobility in polymers are some of the factors determining the brittle transition.
7. Surface treatments such as introduction of compressive stresses at the surface improve the resistance to crack propagation in brittle materials.
8. The majority of failures of structural and machine components are due to fatigue. The resistance to fatigue crack propagation is increased by surface treatments and also by a fine grain size.

PROBLEMS

- 12.1** Show that the critical value of ΔU in Eq. (12.1) is equal to $2\gamma c$. For typical values of γ and c , show that the thermal energy available at most temperatures is unlikely to be of assistance in nucleating a crack of critical dimensions in a crack free material.
- 12.2** Estimate the surface energy of a surface-etched sodium silicate glass, which has a fracture strength of 100 MN m^{-2} and a Young's modulus of 70 GN m^{-2} . Assume that etching has removed all surface cracks. A number of cracks are present inside the glass sample and they vary in length from $1 \text{ }\mu\text{m}$ to $5 \text{ }\mu\text{m}$.
Answer: 0.56 J m^{-2} .
- 12.3** A sheet of glass with a Young's modulus of 70 GN m^{-2} and a surface energy of 0.5 J m^{-2} has an internal crack of length equal to $2 \text{ }\mu\text{m}$. If an external crack of length $10 \text{ }\mu\text{m}$ is introduced by scratching the surface with a diamond tool, by how much would the fracture strength change? Assume that both the cracks are oriented perpendicular to the stress axis.
Answer: Decrease by 102 MN m^{-2} .
- 12.4** A sheet of glass has an internal crack of length equal to $2 \text{ }\mu\text{m}$. What stress should be applied along an axis inclined at 60° to the crack surface to cause fracture? $Y = 70 \text{ GN m}^{-2}$ and $\gamma = 1 \text{ J m}^{-2}$.
Answer: 244 MN m^{-2} .
- 12.5** A tensile stress is applied to a glass rod having a surface crack $2 \text{ }\mu\text{m}$ deep. Find the stress necessary to satisfy the Griffith condition. Check whether this stress is sufficient to cause fracture. Assume the radius of curvature at the tip of the crack to be $2 \text{ }\text{\AA}$. $Y = 70 \text{ GN m}^{-2}$ and $\gamma = 1 \text{ J m}^{-2}$.
Answer: 149 MN m^{-2} .
- 12.6** What is the reason for the variation of strength from specimen to specimen in ceramic materials?
- 12.7** Brittle materials are found to be stronger, if the total volume under test is reduced drastically. Why?
- 12.8** Why are brittle materials used more often in compression than in tension in structural design?
- 12.9** The ends of two cables are to be tied to a ceramic insulating support such that they are not in electrical contact with each other. How would you design the support, such that the ceramic does not break, when the cables are pulled apart?
- 12.10** A sodium chloride crystal is brittle when tested in air but becomes ductile when tested under water. Why is this so?

- 12.11** A heat treated steel chisel and a glass window pane are both brittle. Explain why the chisel is strong and the window pane is weak.
- 12.12** Suggest one of the following alloys for use as a container for liquid oxygen in a rocket, giving your reasons for the choice: mild steel, copper and austenitic stainless steel.
- 12.13** When a sodium silicate glass is immersed in a lithium nitrate bath at 260°C for a few minutes, cracks develop on the surface. Why?
- 12.14** A sample of borosilicate glass contains two cracks: a surface crack of $1\text{ }\mu\text{m}$ deep and an inner crack of $1.8\text{ }\mu\text{m}$ long. Both the cracks are normal to the tensile axis. Determine which one will propagate first on increasing the applied load and at what stress? $Y = 65\text{ GN m}^{-2}$; $\gamma = 0.5\text{ J m}^{-2}$.
Answer: Surface crack at 144 MN m^{-2} .
- 12.15** A sheet of glass, with $Y = 70\text{ GN m}^{-2}$ and $\gamma = 0.5\text{ J m}^{-2}$, has an internal crack of length equal to $2\text{ }\mu\text{m}$. A surface crack of $0.8\text{ }\mu\text{m}$ depth is introduced by scratching the surface with a sharp tool. Adsorption at the surface lowers the surface energy by 50%. Determine which crack will propagate first, on increasing the applied stress and at what value of the stress?
Answer: Surface crack at 118 MN m^{-2} .
- 12.16** Estimate the fracture strength of a sodium silicate glass with $Y = 70\text{ GN m}^{-2}$ and $\gamma = 0.5\text{ J m}^{-2}$. A number of cracks are present both inside and on the surface of the glass, varying in size from 1 to $5\text{ }\mu\text{m}$. How will the fracture strength change, if the glass is dipped in (i) HF solution, and (ii) LiNO_3 solution?
Answer: 67 MN m^{-2} .

MULTIPLE CHOICE QUESTIONS

- In the Griffith equation, the fracture stress is proportional to
 A. \sqrt{c} B. $1/\sqrt{c}$ C. $1/c$ D. $1/2c$
- The fracture strength of borosilicate glass ($Y = 65\text{ GN m}^{-2}$; $\gamma = 0.5\text{ J m}^{-2}$) with a surface crack of $1\text{ }\mu\text{m}$ deep and with the crack faces parallel to the tensile axis is
 A. 144 MN m^{-2} B. 203 MN m^{-2} C. $\sim Y/6$ D. 0
- If the surface crack causing fracture in a brittle material is made twice as deep, the fracture strength will
 A. decrease by a factor of $\sqrt{2}$ B. decrease by a factor of 2
 C. decrease by a factor of 4 D. not change

4. A silicate glass has a relatively low fracture strength, because
 - A. the Young's modulus of glass is low
 - B. the cracks propagate, before Griffith criterion is satisfied
 - C. plastic deformation during crack propagation causes fracture
 - D. the cracks are sharp and propagate as soon as Griffith condition is met
5. In brittle materials with atomically sharp cracks, the stress concentration at the tip of the crack is a factor of
 - A. 2
 - B. 200
 - C. 5000
 - D. 10^6
6. Liberty ships in World War II failed by brittle fracture due to:
 - A. going above the ductile-brittle transition temperature
 - B. going below the ductile-brittle transition temperature
 - C. glass superstructure
 - D. defective riveting
7. The residual stresses in the interior of a tempered glass are
 - A. nil
 - B. tensile
 - C. compressive
 - D. highly compressive
8. An ion-exchange method of strengthening will be effective for a sodium silicate glass if it is dipped in
 - A. LiNO_3
 - B. NaNO_3
 - C. KNO_3
 - D. none of these
9. Tick the methods that improve fatigue resistance of materials
 - A. fine grain size
 - B. shot peening
 - C. polishing the surface
 - D. decarburizing a steel
10. The fatigue strength of mild steel is
 - A. equal to its tensile strength
 - B. more than its tensile strength
 - C. equal to its yield strength
 - D. lower than its yield strength

Answers

- | | | | | |
|------|------|------|------------|-------|
| 1. B | 2. C | 3. A | 4. D | 5. B |
| 6. B | 7. B | 8. C | 9. A, B, C | 10. D |

Suggestions for Further Reading

F.A. McClintock and A.S. Argon (Eds.), *Mechanical Behavior of Materials*, Addison-Wesley, Reading, Mass. (1966), Chaps. 15–17.

A.S. Tetelman and A.J. McEvily, Jr., *Fracture of Structural Materials*, John Wiley, New York (1967).

13 Oxidation and Corrosion

The attack of environment on materials can take different forms. At high temperatures, metals get oxidized. Polymers react with oxygen and degrade. Ceramic refractories may be dissolved by the liquid they are in contact with. Materials used in a nuclear reactor may undergo radiation damage. Metals corrode under atmospheric conditions. Rusting of iron is a common example of atmospheric corrosion.

In this chapter, as examples, we first consider the high temperature oxidation of metals. The rest of the chapter is devoted to the atmospheric corrosion of metals. The importance of the study of corrosion, in these days of scarce raw material resources, can be hardly overemphasized. One estimate puts the loss due to corrosion at 3% of the annual income of countries.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Pilling–Bedworth number	–	–	–
Standard electrode potential V	volt	V	–
Current density J	ampere per square metre	A m^{-2}	mA/cm^2
Electrical conductivity σ	per ohm per metre	$\text{ohm}^{-1} \text{m}^{-1}$	mho/cm

Constants

Faraday's constant $F = 96.49 \text{ kC mol}^{-1}$ (of electrons)

OXIDATION

13.1 Mechanisms of Oxidation

At ordinary temperatures, the stable state of most metals is in the form of an oxide. This is seen from Table 13.1, which lists the free energy of formation ΔG of one mole of different oxides at 25°C. Except in the case of gold, the free energy change is negative, indicating that the stable form is the oxide. Only gold occurs in the metallic form in nature. All the other metals are to be reduced to the metallic state by an extraction process. While using most metals in the metastable state, we depend on the fact that *the thermal energy at the service temperature is small enough to keep the oxidation rate within the desired limits*. As the service temperature increases, the oxidation becomes a serious problem.

TABLE 13.1
Free Energy of Formation of Metal Oxides at 25°C (kJ mol⁻¹)

Al ₂ O ₃	Cr ₂ O ₃	TiO ₂	Fe ₂ O ₃	MgO	NiO	Cu ₂ O	Ag ₂ O	Au ₂ O ₃
-1576	-1045	-853	-740	-568	-217	-145	-13	+163

The rate of oxidation of a metal at an elevated temperature depends on the nature of the oxide layer that forms on the metal surface. For good oxidation resistance, *the oxide layer should be adherent to the surface*. The adherence of an oxide film is dependent on the ratio of the volume of oxide formed to that of metal consumed during oxidation. This ratio known as the *Pilling–Bedworth ratio* as given in Table 13.2 for some oxides. If the ratio is less than unity, tensile stresses will be set up in the oxide layer. The oxide being brittle cannot withstand tensile stresses and, therefore, it cracks and is not protective against further oxidation. If the ratio is more than unity, the oxide layer will be in compression, will uniformly cover the metal surface and be protective. If the ratio is much greater than unity, there is the risk of too much compressive stresses being set up, again resulting in the cracking of the layer.

TABLE 13.2
The Pilling–Bedworth Ratio for Some Oxides

K ₂ O	Na ₂ O	MgO	Al ₂ O ₃	NiO	Cu ₂ O	Cr ₂ O ₃	Fe ₂ O ₃
0.41	0.58	0.79	1.38	1.60	1.71	2.03	2.16

When a metal is subject to alternate heating and cooling cycles in service, the relative thermal expansion of the oxide and the metal also determines the stability of the protective layer. Thermal shock caused by rapid heating or cooling may cause the layer to crack. If the oxide layer is volatile as is the case with molybdenum and tungsten oxides at high temperatures, there will be no protection.

When the oxide layer is adherent to the metal surface, further oxidation can take place only by means of diffusion through the oxide layer of the oxygen anions or the metal cations. When the diffusion of the oxygen anions controls the oxidation rate, oxidation takes place at the metal-oxide interface. If the metal cations diffuse through the oxide layer in the opposite direction, then oxidation takes place at the oxide-oxygen interface. As the oxide layer increases in thickness as a function of time, the diffusion distance through the layer also increases.

Example 13.1 Assuming the oxidation of a metal to be controlled by the diffusion of oxygen anions through the oxide layer, find the functional dependence of oxide thickness x on time t .

Solution Referring to Fig. 13.1, we have a constant oxygen concentration c_2 in the atmosphere and c_1 at the metal-oxide interface, $c_2 > c_1$. The oxidation

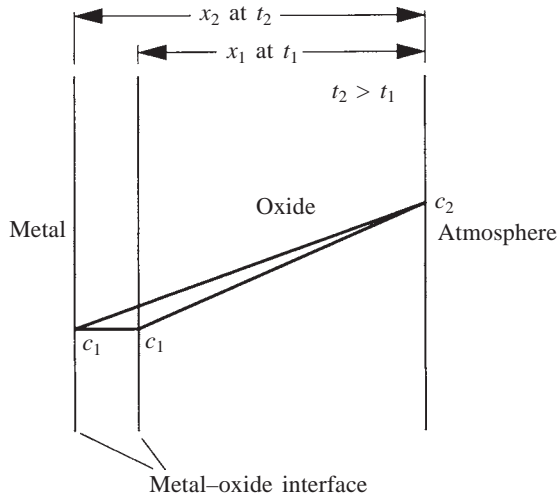


Fig. 13.1 The concentration-distance profile for the diffusion of oxygen anions during oxidation.

takes place at the latter interface. At $t_2 > t_1$, the oxide thickness is more and the concentration gradient across it is less. The rate of thickening of the oxide layer dx/dt is proportional to the flux J so that

$$\frac{dx}{dt} \propto J = -D \frac{dc}{dx} = -D \frac{c_2 - c_1}{x}$$

$$\int x \, dx \propto -D(c_2 - c_1) \int dt$$

$$x^2 \propto Dt$$

$$x \propto \sqrt{Dt}$$

The thickness of the layer increases as the square root of time at constant temperature.

The above square root relationship is called the *parabolic law of oxidation*. Many metals obey the parabolic law at some temperature range of oxidation. Note that the square of the layer thickness is proportional to the diffusion coefficient D . Recall from Chap. 8 that the diffusion coefficient increases with temperature in an exponential manner. Correspondingly, the oxidation rate also increases exponentially with temperature, thus rapidly aggravating the problem of protection as the temperature increases. The activation energy for oxidation is the same as the activation energy for diffusion through the oxide layer. The diffusion coefficient at a given temperature depends on the nature of the oxide layer as well as its defect structure.

If the Pilling–Bedworth ratio is much greater than unity, the oxide layer tends to crack on reaching a critical thickness. This is especially so if the oxidation process occurs at the metal-oxide interface, where the expansion cannot be accommodated as easily as at the oxide-gas interface. When the excess layer beyond the critical thickness peels off, the oxidation rate becomes constant indicating a constant critical oxide thickness at the metal surface.

13.2 Oxidation Resistant Materials

The key to the control of oxidation is to form a *protective layer of oxide* on the metal surface. The oxide layer should also offer a *high diffusion barrier* to the motion of the species that control the oxidation process. The electrical conductivity of the oxide is a measure of the diffusivity of the moving ions, as already discussed in Sec. 8.6. A very low conductivity indicates negligible deviation from stoichiometry and hence a low diffusivity, which effectively reduces the oxidation rate. A high conductivity implies a significant deviation from stoichiometry. A nonstoichiometric oxide will allow much faster diffusion of the diffusing species.

The oxidation resistance of a metal is improved by the addition of suitable alloying elements to the base metal. The alloying element must be present in sufficient concentration to produce the desired oxide layer. The most common alloying elements *added to iron* for this purpose are *chromium, aluminium and nickel*. The oxidation rate of iron as a function of chromium content is shown in Fig. 13.2. The rate decreases with increasing chromium content. The addition of chromium enables the formation of a thin protective layer of Cr_2O_3 on the surface of iron. The electrical conductivity of Cr_2O_3 is one-thousandth of that of Fe_2O_3 . Up to 10% of chromium is alloyed with iron for improving the oxidation resistance of oil refinery components. Alloys with greater than 12% chromium are known as *stainless steels*. They are used for turbine blades, furnace parts and valves for IC engines. 12% chromium gives excellent oxidation resistance up to

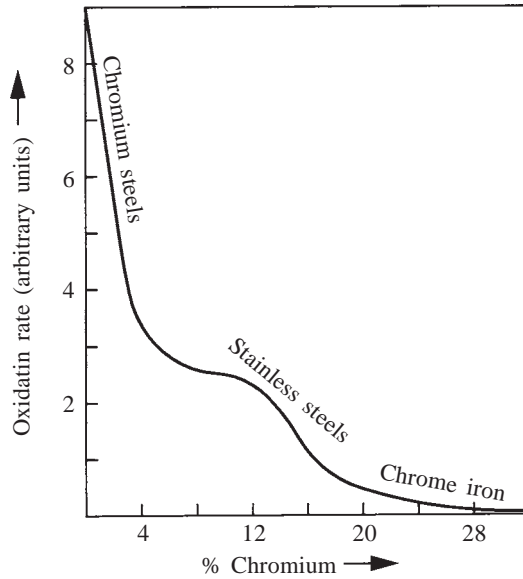


Fig. 13.2 The oxidation rate of iron decreases with increasing chromium content.

1000°C, while 17% chromium is used above 1000°C. 18-8 stainless steel, which contains 18% Cr and 8% Ni, is among the best commercially-available oxidation resistant alloys. 24% Cr, 5.5% Al and 2% Co alloyed with iron, known as *kanthal*, is used for furnace windings up to 1300°C.

80% Ni–20% Cr (*nichrome*) and 76% Ni, 16% Cr and 7% Fe (*inconel*) have excellent oxidation resistance and good mechanical properties. 10% Cr alloyed with nickel (*chromel*) and 2% Al, 2% Mn and 1% Si alloyed with nickel (*alumel*) are used up to 1100°C as heat resistant thermocouple wires. Molybdenum in a protective atmosphere of hydrogen can be used for furnace windings up to 1500°C.

Aluminium is normally covered with a highly protective oxide film. Therefore, there is usually no need to add alloying elements to aluminium to improve its oxidation resistance. Similarly, titanium forms a protective oxide layer. Addition of aluminium, beryllium or magnesium to copper improves its oxidation resistance.

CORROSION

The direct losses due to corrosion of structural and machine components is estimated to be 30 billion dollars annually. It is, therefore, essential for an engineer to understand the basic principles of corrosion and the methods of protection against corrosion.

13.3 The Principles of Corrosion

13.3.1 The Electrode Potential

If a piece of iron rod (electrode) is immersed in a polar solvent like water, some of the metal ions leave the crystal and go into solution. They get hydrated, that is, they form bonds with the polar water molecules. As the metal continues to dissolve, more and more electrons are left back and a net negative charge builds up in the metal. Similarly, the presence of excess positive ions builds up a positive charge in the adjacent solution. These opposing electrical layers discourage further dissolution of the metal. Soon, a dynamic equilibrium is reached with no net flow of metal ions into the solution.

The potential developed by an electrode in equilibrium is a property of the metal forming the electrode. This potential is measured under standard conditions with a pure metal as the electrode and an electrolyte containing unit concentration of ions of the same metal. As only potential differences can be measured in an experiment, the potential of an electrode is determined against a *standard hydrogen electrode* whose potential is taken to be zero. The standard potentials for various metals are listed in Table 13.3. The sign convention used here is that adopted by the International Union of Pure and Applied Chemistry.

TABLE 13.3
Standard Electrode Potentials of Metals

	Electrode system	Standard potential V_0 at 25°C, volt	Electrode system	Standard potential V_0 at 25°C, volt	
<i>Noble end</i>	Au/Au ³⁺	+1.50	Zn/Zn ²⁺	-0.76	
	Ag/Ag ⁺	+0.80	Mn/Mn ²⁺	-1.18	
	Cu/Cu ²⁺	+0.34	Zr/Zr ⁴⁺	-1.53	
	H ₂ /H ⁺	0.0	Ti/Ti ²⁺	-1.63	
	Pb/Pb ²⁺	-0.13	Al/Al ³⁺	-1.66	
	Sn/Sn ²⁺	-0.14	U/U ³⁺	-1.80	
	Mo/Mo ³⁺	-0.20	Be/Be ²⁺	-1.85	
	Ni/Ni ²⁺	-0.25	Mg/Mg ²⁺	-2.37	
	Co/Co ²⁺	-0.28	Na/Na ⁺	-2.71	
	Cd/Cd ²⁺	-0.40	Ca/Ca ²⁺	-2.87	
	Fe/Fe ²⁺	-0.44	K/K ⁺	-2.93	
	Cr/Cr ³⁺	-0.74	Li/Li ⁺	-3.05	<i>Active end</i>

Gold at the top of the list is the most noble metal and will not dissolve easily. Lithium at the bottom of the list is the most active and base metal; it will go into solution readily.

The standard potential will be changed if the metal is not pure or if the metal ion concentration in the electrolyte is not unity. The potential V under nonstandard conditions is given by

$$V = V_0 - \frac{RT}{nF} \ln \left(\frac{M}{M^+} \right) \tag{13.1}$$

where V_0 is the standard potential, M^+ is the metal ion concentration in the electrolyte, M is the concentration of the metal in the electrode, n is the valence of the metal ion and F is Faraday constant equal to 96.49 kC/mole of electrons. Under standard conditions, the second term on the right side of Eq. (13.1) is zero. If $M^+ < 1$, that is, if the electrolyte is deficient in metal ions, the potential decreases and goes more towards the active end of Table 13.3. If the electrode is in the alloyed condition, $M < 1$, and then the potential increases and goes towards the noble end.

13.3.2 The Galvanic Series

The structure and composition of the alloys used in service are complex. The environmental conditions which provide the electrolyte are also difficult to define in terms of M^+ . The role of environment in determining the corrosion rate of mild steel is seen from the data in Table 13.4.

TABLE 13.4
Role of Environment on the Corrosion Rate of Mild Steel

Environment	Corrosion rate, mm per year
Dry and nonindustrial	0.001
Marine, humid and nonindustrial	0.02
Humid and industrial	0.2

To describe the tendency to corrode in a given environment, common metals and alloys are arranged on a qualitative scale called the *galvanic series*. For example, the galvanic series for the sea water environment is given in Table 13.5. In sea water, 18-8 stainless steel in the passive condition at the top of the table is the least active and magnesium at the bottom of the table is the most active.

TABLE 13.5
Galvanic Series in Sea Water

<i>Noble end</i>	18-8 stainless steel (passive)
	Nickel (passive)
	Copper
	Brass
	Tin
	Lead
	18-8 stainless steel (active)
	Mild steel
	Alclad
	Aluminium
	Zinc
<i>Active end</i>	Magnesium

13.3.3 The Galvanic Cells

Figure 13.3 shows a galvanic cell with zinc and copper as the two electrodes. The excess electrons left back at the zinc electrode are more than that in the

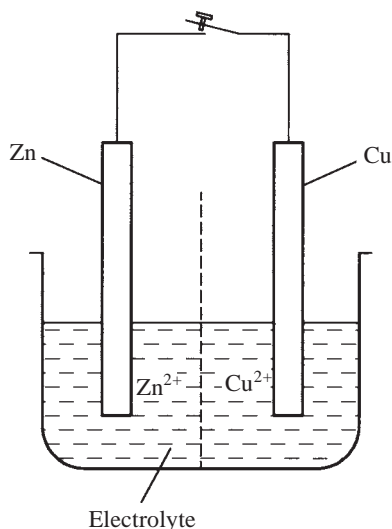


Fig. 13.3 The zinc-copper galvanic cell. On closing the circuit, the zinc anode starts to corrode.

copper electrode, as zinc is less noble than copper, see Table 13.3. If the electrodes are pure and if the metal ion concentration near each electrode is unity, the cell will develop a potential of $0.34 - (-0.76) = 1.10$ V. On closing the circuit, the excess electrons at the zinc will flow through the external circuit to the copper electrode under the concentration gradient. The conventional current will flow in the opposite direction.

If the current continues to flow, the zinc electrode will gradually dissolve in the electrolyte or in other words corrode. Zinc is the *anode* of the galvanic cell. The anode reaction is the dissolution of zinc:



Copper is the cathode of the galvanic cell. Copper will deposit on the cathode through the following reduction reaction:



If cupric ions are not available in the electrolyte, the reduction at the cathode may still proceed through other reactions. Hydrogen ion may get reduced, resulting in the evolution of hydrogen gas at the cathode:



Alternatively, the reduction may take place through oxygen and water combining with the excess electrons to produce hydroxyl ions:



Two different phases in the same metal may form a galvanic couple at the *microstructural level*. For example, cementite in steel is more noble (cathodic) with respect to ferrite. Under corrosive conditions, ferrite will form the anode and corrode away. The finer is the distribution of ferrite and cementite in a steel, the more are the number of galvanic cells that can form and the faster will be the corrosion rate. This effect is illustrated in Fig. 13.4, where the corrosion rate

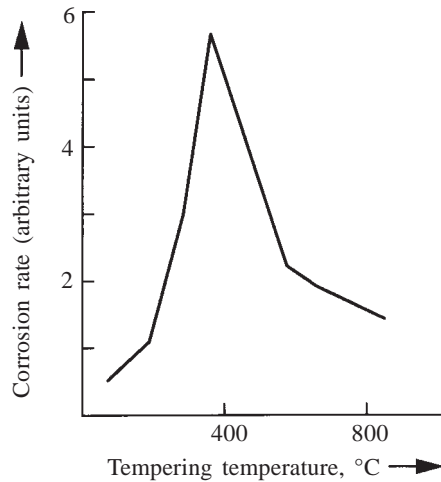


Fig. 13.4 Corrosion rate of a quenched and tempered steel as a function of the tempering temperature.

of a quenched and tempered steel is plotted against the tempering temperature. Initially, the quenched steel has only one phase, martensite. On tempering the martensite, cementite precipitates in a fine form, resulting in the formation of a number of galvanic cells, with cementite and ferrite as the two 'electrodes'. Accordingly, the corrosion rate increases initially with increasing tempering temperature, see Fig. 13.4. When the tempering temperature is above 400°C, the corrosion rate decreases. Here, during tempering, the cementite particles coalesce or coarsen into larger particles resulting in a decrease of the number of galvanic cells, as compared to that at a lower tempering temperature. Hence, a decrease in the corrosion rate is observed.

A galvanic cell may be set up due to differences in concentration of the metal ion in the electrolyte. If, in a region of the electrolyte, the metal ions are deficient, the metal near that region will be anodic with respect to the metal near a different region where the electrolyte has excess metal ions. This can be easily deduced from Eq. (13.1). A galvanic cell set up due to such differences in metal ion concentration is known as a *concentration cell*.

A concentration cell can also arise due to *differences in oxygen concentration*. The cathodic reaction (13.5) takes place more readily where oxygen is available, so that an oxygen-rich region is cathodic with respect to an oxygen-depleted region. This results in *crevice corrosion*, which occurs at inaccessible locations (crevices) deficient in oxygen. Examples of such locations are the interfaces of two coupled pipes, threaded connections and areas covered with rust or dirt. Corrosion occurs just below the water line in a tank. This location being deficient in oxygen is anodic to the metal just above the water line. An underground pipeline that goes through impervious clays in some regions and through porous sands in some other regions may corrode in the clay region.

A galvanic cell can form due to different residual stresses in the same metal. The stressed region is more active and is anodic with respect to a stress-free region. Such *stress cells* can form between regions of different dislocation density in a cold-worked metal or in a polycrystalline metal, where the grain boundaries are anodic to the interior of the grains. A bent wire is likely to corrode at the bend, where it has been plastically deformed.

13.3.4 Polarization

When a galvanic cell is short circuited, a corrosion current flows through the cell. This current would set up differences in concentration of the metal ions near the electrodes. In the zinc–copper cell, due to continued dissolution of zinc at the anode, the zinc ion concentration near the anode builds up. Unless the diffusion rate of these ions is fast enough to keep the electrolyte composition uniform, the potential of the zinc electrode will shift towards the noble end, that is, towards that of copper. Similarly, at the cathode, the cupric ions are being used up in the cathode reaction, resulting in a deficiency of these ions near the cathode. This shifts the cathode potential towards the active side, that is, towards that of zinc. In addition, the two electrode potentials tend to move towards each other due to the limited rates of the anodic and the cathodic reactions (which proceed by the reacting species crossing an activation barrier). Figure 13.5 shows schematically

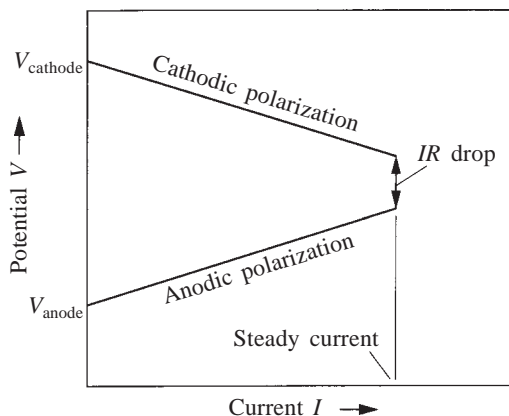


Fig. 13.5 The anode and the cathode polarization as a function of current I .

the variations in the cathode and the anode potentials with increasing current density. The electrodes are said to be *polarized*, as the current flows. A steady state current is established, when the potential difference between the cathode and the anode becomes equal to the IR drop through the electrolyte.

13.3.5 Passivation

A piece of iron or steel readily dissolves in dilute nitric acid, but may become resistant or passive in concentrated nitric acid. The concentrated acid oxidizes the iron effectively and produces a thin protective layer on the surface. Dilute acid is not strong enough to oxidize and hence continues to attack. The formation of an oxide film on a corroding metal can passivate it in a similar fashion.

On increasing the potential of a metal electrode, the current density increases at first. When the current density reaches a critical value, it may abruptly fall to a much lower value and remain more or less constant for some further increase in potential. This phenomenon is called *passivation*. For chromium, the critical current density just before passivation is about 200 A m^{-2} . The current after passivation is less than 0.1 A m^{-2} . This big drop in current density is associated with the simultaneous formation of a thin oxide layer on the metal surface. The phenomenon of passivation affords an important means of corrosion prevention.

13.4 Protection against Corrosion

The methods of corrosion prevention are based on the above principles of corrosion. From the table of electrode potentials, one can deduce that use of *noble metals* will prevent corrosion. However, it is clear that the choice of a material is dependent on many other factors. We can use noble metals only in very limited applications such as ornaments and delicate scientific instruments.

Alternatively, we can design to *avoid physical contact between dissimilar metals* so that a galvanic couple does not form. Yet, we often come across designs such as a steel screw in brass marine equipment, a Pb–Sn solder on copper wire and a steel shaft in a bronze bearing. When contact between dissimilar metals is unavoidable, it is necessary to see that the metal forming the *anode does not have a small surface area* as compared to the cathode. As the same corrosion current passes through the anode and the cathode, a small anode area would mean a high current density at the anode and a consequent high rate of corrosion. A copper nut and bolt is permissible on a large steel plate or pipe, but not a steel bolt and nut on a copper article of large surface area.

In materials with two-phase structures, where the phases form galvanic cells at the microstructural level, the contact between the two phases is inevitable. We have already referred to the example of untempered, tempered and overtempered martensites, which corrode at different rates (see Fig. 13.4). The two-phase structure of duralumin corrodes faster than pure aluminium. However, the optimum aged two-phase structure is often required for obtaining the desired mechanical properties. In the case of duralumin, the corrosion prevention is

achieved by making *Alclad*, which has two thin pure aluminium sheets covering either side of a duralumin sheet.

Removing a cathodic reactant such as dissolved oxygen from water by means of a chemical reagent may prevent corrosion. For example, sulphites are used for this purpose in boiler feed water and cooling water systems. *Inhibitors* form a protective layer on the metal surface and prevent corrosion. Anodic inhibitors are oxidizing anions such as nitrites and chromates. For example, nitrites promote the formation of a thin passivating oxide film on iron. These must be used in sufficient quantity to cover the whole surface. Otherwise, small uncovered areas may lead to severe localized corrosion known as *pitting*. Cathodic inhibitors generally form a thick protective film on the surface.

Vapour phase inhibitors consist of nitrite or bicarbonate anions attached to a heavy organic cation. The inhibitor compound is placed alongside the metal part to be protected in the storage room. The compound has a vapour pressure of about 10^{-3} atm at ambient temperature, so that it evaporates rapidly to ensure its adequate availability in the vicinity of the metal surface. Sewing needles are first wrapped in thin paper saturated with a vapour phase inhibitor (ethanolamine acetate) and then in thick black paper, which retains the vapour.

Metallic coatings are used for corrosion prevention. If a metal coating is noble with respect to the underlying metal, it is necessary to avoid flaws in the coating such as cracks and pores. Such flaws could initiate corrosion, with the coating acting as the cathode and the underlying metal as the anode. As the exposed part of the anode at a crack or pore is very small, corrosion takes the form of a localized attack. Tin on steel is an example, exhibiting localized *pinhole* attack.

The tin coating on a steel article is produced either by dipping it in molten tin or by electroplating. The most common use of *tinplate* is for making food containers. The pinhole corrosion referred to above can take place on the outside of the container. Inside, many organic acids that are present in foodstuffs and fruit juices form complexes with tin. The concentration of the stannous ions is thereby lowered. The potential of tin decreases enough to make it anodic with respect to iron. In the absence of a suitable cathodic reaction, the corrosion rate is low.

If the metal coating is baser than the substrate (e.g., zinc or aluminium on steel), *galvanic protection* is offered to the underlying metal. The coating is anodic and corrodes first. Zn and Al, however, become passive after the initial attack. *Galvanized iron* (G.I.) is produced by dipping a low carbon steel sheet in molten zinc bath at about 450°C. On cooling in air, the zinc coating crystallizes forming 'zinc flowers'. Articles such as buckets and drums made of galvanized iron are very suitable for aqueous environments. In the presence of oxygen, zinc hydroxide is precipitated as a protective layer. Aluminium coatings are deposited by the process of *calorizing* mainly to improve the oxidation resistance of steel. Decorative chromium plating is done over a first coating of nickel on automobile exterior components. Hard chromium plating produces a thicker, wear resistant surface.

Nonmetallic coatings such as enamel, oil, paint and tar act by simply excluding water and oxygen and by providing a layer of high electrical resistance.

For articles used indoors, it is often enough to coat them with one layer of paint. For the protection of outdoor structures such as bridges, elevator cranes, harbour equipment, railway carriages, automobiles and bicycles, which are exposed to the action of the atmosphere, paints are of more complex composition. Also, their method of application is more intricate, involving several coatings.

Cathodic protection is based on the electrochemical principles already outlined. Protection by galvanic action is provided by connecting the structure to be protected to a *sacrificial anode*, which is periodically replaced. The objects thus protected are either buried in the earth or immersed in water, e.g., underground and underwater pipelines, foundations, piers, jetties, ships and offshore structures. A piece of magnesium, zinc or aluminium is attached to the article to be protected making electrical contact with it. In the artificial galvanic cell thus created, the added electrode piece is the anode, which is gradually destroyed.

Example 13.2 For cathodic protection, a ship hull requires a current density of 15 mA m^{-2} . Magnesium is used as the sacrificial anode. What is the quantity of magnesium required per square metre of the hull surface for a design life of ten years?

Solution One mole of magnesium (0.0243 kg) will be used up as anode for every two Faraday of charge, $2 \times 96490 \text{ C (A s)}$, as magnesium is divalent. The charge required per m^2 of hull surface for a design life of 10 yr = $15 \times 10^{-3} \times 10 \times 365 \times 24 \times 3600 \text{ A s}$. So,

the amount of magnesium needed

$$= \frac{0.0243 \times 15 \times 10^{-3} \times 10 \times 365 \times 24 \times 3600}{2 \times 96490}$$

$$= 0.6 \text{ kg m}^{-2}$$

Cathodic protection can also be obtained by using an external dc source in conjunction with an inert (or expendable) electrode. The negative end of the dc source is connected to the metal to be protected (cathode) and the positive end to the inert anode. Thus, the impressed voltage ensures that the potential of the part to be protected is such as to keep it cathodic. Inert anodes include platinum in the form of a thin electrodeposit on titanium and stainless steel. Expendable electrodes are usually scrap of the same material that is being protected.

Alloying additions such as chromium to iron (as in stainless steel) improve the corrosion resistance by reducing the critical current density for passivation. Steels with more than 12% chromium require only a very small critical current density, in order to become passive.

An improperly heat treated 18/8 stainless steel may exhibit *intergranular corrosion*. Chromium carbide may precipitate at the grain boundaries, due to excessive residual carbon in the stainless steel or due to prolonged exposure to

an elevated temperature, as shown in Fig. 13.6. The region immediately next to the chromium carbide particles is denuded of chromium and, therefore, becomes

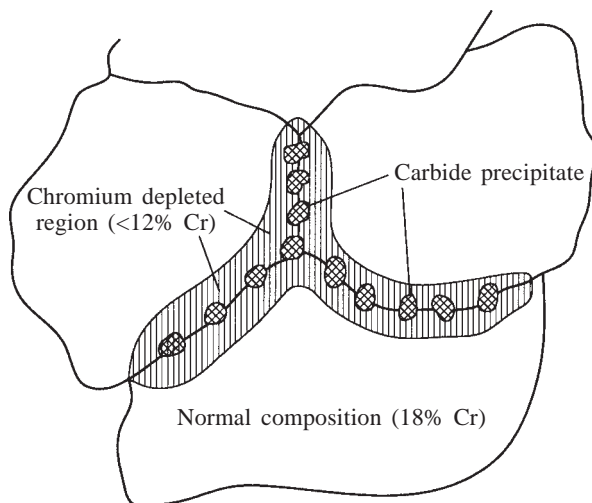


Fig. 13.6 Intergranular corrosion of 18-8 stainless steel.

anodic to the interior of the grain. A galvanic cell is thus set up and corrosion occurs close to the grain boundaries. Intergranular corrosion can be prevented in three ways:

- (i) by *reducing the carbon level* in the stainless steel to less than 0.05%;
- (ii) by *quenching* the stainless steel from a high enough temperature to prevent chromium carbide precipitation; and
- (iii) by *adding strong carbide forming elements* to the steel such as Nb or Ti, so that carbon precipitates as niobium carbide or titanium carbide and *not* as chromium carbide.

Some special types of corrosion are also known. One such type is the *dezincification*. During corrosion of alpha brass, the zinc is preferentially dissolved from the brass, leaving a spongy mass of copper of little strength. Dezincification is prevented by adding 0.04% arsenic to brass. The phenomenon of *stress corrosion* occurs under the combined action of a corrosive environment and a mechanical stress. In *season cracking*, a cold-worked brass with high residual stresses cracks along the grain boundaries in environments containing ammonia. The grain boundaries have piled-up dislocations in their neighbourhood due to the cold working. The high energy due to this causes the boundaries to become anodic with respect to the grain interior and selective dissolution of the boundary region occurs. If the cold-worked brass is *stress relieved* at 300°C, season cracking can be avoided. Stress corrosion of *austenitic stainless steels* occurs in *chloride environments*. The crack here is *transgranular*, that is, it propagates across the grains. Addition of molybdenum or increasing the nickel content of the stainless steel reduces the tendency for this type of corrosion. *Caustic embrittlement* of boilers occurs under the combined action of stress and a high concentration of *hydroxyl ions* in the environment. Addition of tannins and phosphates to the environment prevents caustic cracking.

SUMMARY

1. The formation of an adherent oxide film is necessary to improve the oxidation resistance of a metal. The oxide film should offer a high diffusion barrier to the diffusing species that control the oxidation.
2. The common alloying elements added to steel for improving its oxidation resistance are chromium, aluminium and nickel.
3. The standard electrode potential measures the tendency of a metal to go into solution. Metals with positive potentials are noble, as compared to metals with negative potentials.
4. A galvanic cell may form under different conditions, e.g., due to a two-phase structure, due to concentration differences in the electrolyte or due to stress differences.
5. Passivation refers to the formation of a thin, high resistance oxide layer, which reduces the corrosion current by two or three orders of magnitude.
6. (a) Designing to avoid contact between dissimilar metals, (b) use of inhibitors, (c) metallic and nonmetallic coatings, (d) cathodic protection with a sacrificial anode or with an impressed voltage, and (e) passivation by alloy additions are the main methods of corrosion prevention.

PROBLEMS

- 13.1** Calculate the Pilling–Bedworth ratio for CaO. The density of CaO and Ca are 4030 and 1550 kg m⁻³, respectively.
Answer: 0.54
- 13.2** The electrical conductivity of Al₂O₃ is about 10⁻⁵ ohm⁻¹ m⁻¹ and that of Cu₂O is about 10³ ohm⁻¹ m⁻¹. Explain why aluminium can be used to improve the oxidation resistance of copper, but not copper to improve the oxidation resistance of aluminium.
- 13.3** In metals with the Pilling–Bedworth ratio less than unity, the oxidation rate is linear from the beginning of the oxidation process, whereas in metals with ratio substantially larger than unity, the linear rate is observed after an initial parabolic region of oxidation. Explain why this is so.
- 13.4** The activation energy for cuprous ion diffusion in Cu₂O is 160 kJ mol⁻¹. Cuprous ion diffusion controls the oxidation of copper above 500°C. Find the ratio of oxidation rates of copper at 900 and 600°C.
Answer: 280.5.

- 13.5** What is the potential difference that is observed, when a standard zinc electrode is connected to the calomel reference electrode whose potential is + 0.334 V?
Answer: 1.094 V.
- 13.6** Calculate the ampere-hour required to plate out one mole of copper from a copper sulphate solution.
Answer: 53.61 A hr.
- 13.7** Calculate the cupric ion concentration in the electrolyte that is required to make the potential of copper to become equal to that of the standard zinc electrode. Is this potential likely to be realized in practice? Explain.
Answer: 1.0×10^{-37} .
- 13.8** 1 mm thick layer of nickel is to be plated on a steel article. What is the current density required to do this in one hour?
Answer: 8130 A m^{-2} .
- 13.9** Why is tin preferred to zinc on a steel food can? Why is enamelling preferred to paint on a steel refrigerator body?
- 13.10** Show how rust, $\text{Fe}(\text{OH})_3$, can form in a galvanic cell, where iron is the anode and the cathode reaction is as given in Eq. (13.5).
- 13.11** Which of the following coatings is expected to protect iron even if the coating has cracks: Polyethylene, enamel, Zn, Ni, Sn and Pb? Give a reason for your answer.
Answer: Zn.

MULTIPLE CHOICE QUESTIONS

- For a protective oxide layer to form, the ratio of the volume of oxide formed to that of metal consumed should be
 A. < 1 B. 1 C. > 1 D. $\gg 1$
- The keys to good oxidation resistance of an oxide film are
 A. low electrical conductivity B. low electrical resistivity
 C. continuous oxide film D. porous oxide film
- In parabolic rate of oxidation, the oxide thickness is proportional to
 A. $t^{1/2}$ B. t C. t^2 D. none of these
- Tick the elements added to iron to improve its oxidation resistance
 A. zinc B. chromium C. magnesium D. aluminium
- If the activation energy for oxidation is 100 kJ/mol, the ratio of oxidation rates at 800 and 500°C is
 A. 8270 B. 78 C. 1.0 D. 0.031

6. The standard potential of a zinc electrode is -0.76 V, its potential at 300 K at a zinc ion concentration of 0.001 in the electrolyte is ($F = 96490$ C/mol)
 - A. -0.85 V B. -0.67 V C. -0.72 V D. 0 V
7. The potential of a galvanic cell of copper (potential of $+0.34$ V) and aluminium (potential of -1.66 V) is
 - A. 2.00 V B. -1.32 V C. 1.32 V D. 0 V
8. The peak in corrosion rate of martensite occurs when
 - A. in untempered condition B. tempered at 800°C
 - C. tempered at about 400°C D. in spherodized state
9. Pore-free coating is required when
 - A. coating is noble with respect to the protected metal
 - B. coating is base with respect to the protected metal
 - C. coating has the same potential as the protected metal
 - D. none of these
10. For cathodic protection at a current density of 10 mA m^{-2} , the quantity of zinc (atomic weight = 65.4) required per m^2 of ship hull per year is ($F = 96490$ C/mol)
 - A. 0.1 kg B. 0.2 kg C. 107 kg D. 213 kg

Answers

- | | | | | |
|------|---------|------|---------|-------|
| 1. C | 2. A, C | 3. A | 4. B, D | 5. B |
| 6. A | 7. A | 8. C | 9. A | 10. A |

Source for Experimental Data

- S. Lamb (Ed.) *Practical Handbook of Stainless Steels and Nickel Alloys*, ASM International, Materials Park, Ohio (1999).

Suggestions for Further Reading

- D.A. Jones, *Principles and Prevention of Corrosion*, Prentice-Hall, Englewood-Cliffs, New Jersey (1996).
- J.C. Scully, *The Fundamentals of Corrosion*, Pergamon, Oxford (1975).

14 Conductors and Resistors

The electrons in the outermost orbitals of the atoms of a solid determine its electrical properties. In this chapter, the free electron model is introduced, starting with a brief description of the wide range of resistivity of materials. The justification for Ohm's law is given, in terms of the free electron motion and the collision processes in crystals. Materials for conductive and resistive functions are discussed. The phenomenon of superconductivity and the recent discovery of oxide superconductors are discussed at the end.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
Resistivity ρ	ohm metre	ohm m	microohm-inch, ohm-cm
Temperature coefficient of resistance α	per kelvin	K^{-1}	per $^{\circ}F$
Conductivity σ	per ohm per metre	$ohm^{-1} m^{-1}$	mho/cm
de Broglie wavelength λ	metre	m	\AA
Wave number k	per metre	m^{-1}	—
Kinetic energy E } Fermi energy E_F }	joule	J	erg, eV
Drift velocity \bar{v}	metre per second	$m s^{-1}$	—
Field gradient ε	volt per metre	$V m^{-1}$	volts/mil
Current density J_e	ampere per m^2	$A m^{-2}$	amp/cm ²

Constants

Planck's constant h	$= 6.626 \times 10^{-34} \text{ J s}$
Rest mass of electron m_0	$= 9.109 \times 10^{-31} \text{ kg}$
Charge of electron e	$= 1.602 \times 10^{-19} \text{ C}$

14.1 The Resistivity Range

Electrical resistivity (or conductivity) is probably the most remarkable of all physical properties, in that *it varies over 25 orders of magnitude*. To get a feel for this wide range, Table 14.1 lists the electrical resistivity, at room temperature, of a number of materials which are important from the engineering point of view. The materials fall into three broad categories.

TABLE 14.1
The Resistivity of Materials (ohm m)

10^{-9}	10^{-7}	10^{-5}	10^{-3}	10^{-1}	10^1	10^3
Ag						
Cu Al	Ni	Sb Bi	Ge	Ge		Si
Au	Fe	Graphite	(doped)	(pure)		(pure)
← Metals →			← Semiconductors →			
10^5	10^7	10^9	10^{11}	10^{13}	10^{15}	10^{17}
Window glass		Bakelite	Porcelain	Lucite	PVC	SiO ₂
			Diamond	Mica		(pure)
			Rubber, Nylon			
			Polyethylene			
← Insulators →						

Conductors are metals and alloys. Gold, silver and copper are among the best conductors of electricity. Therefore, their electrical resistivities are the lowest, as shown in Table 14.1. They are followed by aluminium whose resistivity is 60% higher than that of copper. Transition metals such as iron and nickel are not as good conductors as the above. Still poorer conductors are the semimetals of the fifth column, e.g., antimony and bismuth. Graphite, with one of its bonding electrons resonating between the (sp^2) bonds, also fall in this category of semimetals. The electrical resistivity of conductors ranges from 10^{-9} to 10^{-3} ohm m. The electrical conductivity, being the reciprocal of resistivity, ranges from 10^9 ohm⁻¹ m⁻¹ to 10^3 ohm⁻¹ m⁻¹.

When the resistivity is in the range 10^{-3} – 10^3 ohm m, we have the second category of materials known as *semiconductors*. They form the base for a number of solid state devices. Here, the resistivity is a *strong function of small concentrations of impurities*. Doped germanium, with an impurity content of a few tens per million, can have a resistivity about two orders of magnitude lower than that of pure germanium, see Table 14.1. Pure silicon has a higher resistivity than pure germanium.

The third category of materials are *insulators*. Common electrical insulating materials such as polyethylene, bakelite, lucite, mica, PVC, rubber and porcelain fall in this category. The resistivity range for this category extends from 10^4 to beyond 10^{17} ohm m. Here, a difference in resistivity of some twelve orders of

magnitude is noticeable between a silica glass and soda-lime-silica (window glass). This striking difference is a result of the ionic conductivity of window glass. The relatively loosely-bound sodium and calcium cations in soda-lime glass diffuse and conduct much more readily, as compared to the tightly-bound immobile silicon cations in pure silica. Ionic conduction and ionic diffusivity are closely related phenomena, recall Sec. 8.6.

14.2 The Free Electron Theory

The conducting properties of a solid are not a function of the total number of electrons in the solid, as only the *outermost electrons* of the atoms can take part in conduction. In the free electron model, the outermost electrons of an atom are not bound to that atom, but are *free to move through the whole solid*. These electrons have been variously called the free electron cloud, the free electron gas or the Fermi gas.

In the free electron theory, the basic assumption is that *the potential field due to the ion cores is uniform* throughout the solid. The free electrons have the same potential energy everywhere in the solid. Due to the electrostatic attraction between a free electron and the ion cores, this potential energy will be a finite negative value. As we are interested only in energy differences, we can assume this constant potential to be zero. Then the only energy that we have to consider is the *kinetic energy*. This kinetic energy is substantially lower than that of the bound electrons in an isolated atom, as the field of motion for the free electron is considerably enlarged in the solid as compared to the field around an isolated atom. This effect is explained below.

Electrons have both particle-like and wave-like characteristics. The de Broglie wavelength of an electron λ is related to its momentum mv as

$$\lambda = \frac{h}{mv} \quad (14.1)$$

where h is Planck's constant, m is the mass of the free electron and v is its velocity. The wavelength is inversely related to the magnitude of the wave number vector \mathbf{k} :

$$k = \frac{2\pi}{\lambda} \quad (14.2)$$

As the velocity of the free electrons is much smaller than that of light, we can ignore relativistic effects and use the classical relation for kinetic energy E :

$$E = \frac{1}{2}mv^2 \quad (14.3)$$

Substituting Eqs. (14.1) and (14.2) into Eq. (14.3), we obtain

$$E = \frac{h^2k^2}{8\pi^2m} \quad (14.4)$$

The kinetic energy E increases as the square of the wave number. This parabolic relationship between E and k is shown in Fig. 14.1. As λ is inversely related to k , the electron with the largest de Broglie wavelength will have the lowest kinetic energy. With the enlarged field of motion in the solid, the electrons can have larger wavelengths and hence lower kinetic energies.

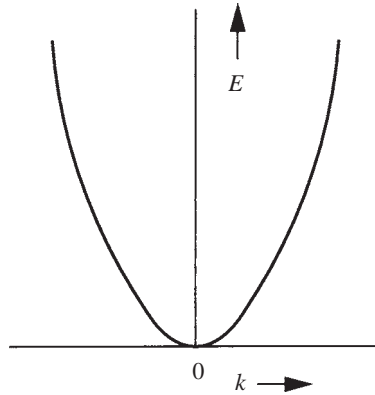


Fig. 14.1 The parabolic relationship between the kinetic energy E of a free electron and its wave number k .

Even though the variation in the kinetic energy E is shown to be continuous in Fig. 14.1, these are in fact *very closely spaced discrete energy levels*, as the quantum restrictions arising from the Pauli exclusion principle apply to the free electrons as well. Let us consider a solid, in which the electron motion is unidirectional. Let L be the length of this axis of motion. The longest wavelength that is permissible for the free electron is equal to twice this length as shown in Fig. 14.2, as the amplitude of the wave should be zero *at both ends*

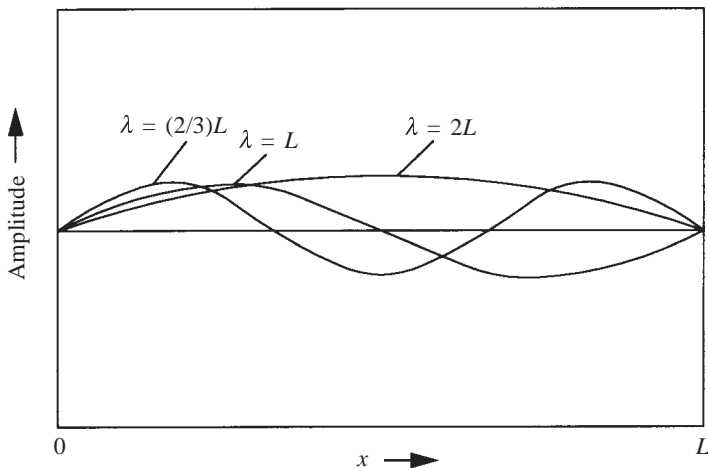


Fig. 14.2 The de Broglie wavelengths of the first few electrons moving along x .

of this length. The next smaller wavelength that is permissible is equal to L , the third one is equal to $(2/3)L$, and so on. The corresponding values of the wavenumber k are π/L , $2\pi/L$, ..., as obtained from Eq. (14.2).

In a solid that is electrically neutral, a free electron having a certain speed in one direction will always be associated with another electron having the same speed but moving in the opposite direction. (If this were not so, there will be a net flow of electrons in one direction, even in the absence of an externally applied field.) In other words, the wave number k takes on both positive and negative values. Thus, for the first two states, we have k equal to $\pm\pi/L$ and $\pm 2\pi/L$. In general, $k = \pm n\pi/L$, where the quantum number n takes on successively increasing integer values. Substituting $k = \pm n\pi/L$ in Eq. (14.4), we obtain

$$E = \frac{h^2 n^2}{8mL^2} \quad (14.5)$$

In a three-dimensional solid, the free electron can move in any direction in space. So, three quantum numbers n_x , n_y and n_z corresponding to the three coordinate axes are used to define each quantum state. n_x , n_y and n_z take on successively-increasing integer values. $(n_x^2 + n_y^2 + n_z^2)$ is substituted in place of n^2 used for the unidirectional motion. Equation (14.5) can be rewritten for three-dimensional motion as

$$E = h^2 \left(\frac{n_x^2 + n_y^2 + n_z^2}{8mL^2} \right) \quad (14.6)$$

Each distinct combination of n_x , n_y and n_z corresponds to a quantum state, where two electrons of opposite spins can reside. Several combinations of n_x , n_y and n_z can result in the same value of E . Quantum states with the same energy are said to be *degenerate*. The successive energy levels calculated from Eq. (14.6) are so close to one another that E can be considered to be varying continuously as illustrated for unidirectional motion in Fig. 14.1.

Example 14.1 Calculate the energy difference between the $n_x = n_y = n_z = 1$ level and the next higher energy level for free electrons in a solid cube of $10 \text{ mm} \times 10 \text{ mm} \times 10 \text{ mm}$.

Solution

$$L = 10 \text{ mm} = 10^{-2} \text{ m}$$

From Eq. (14.6),

$$\begin{aligned} E &= \frac{(6.626 \times 10^{-34})^2 (1^2 + 1^2 + 1^2)}{8 \times 9.109 \times 10^{-31} \times (10^{-2})^2} \\ &= 1.81 \times 10^{-33} \text{ J} \end{aligned}$$

There are many equal energy quantum states above this energy level, with values of n_x , n_y and n_z as $(1,1,2)$, $(1,2,1)$, $(2,1,1)$, For these states,

$$\begin{aligned}
 E &= \frac{(6.626 \times 10^{-34})^2 (1^2 + 1^2 + 2^2)}{8 \times 9.109 \times 10^{-31} \times (10^{-2})^2} \\
 &= 3.62 \times 10^{-33} \text{ J}
 \end{aligned}$$

The energy difference between the first and the next higher energy levels is extremely small, only 1.81×10^{-33} J, so that we are justified in assuming E to be continuously varying with k . Compare this energy difference with the value of thermal energy at temperature as low as 1 K, 1.38×10^{-23} J, which is ten orders of magnitude larger.

The wave forms for the free electrons starting from the lowest energy level are sketched in Fig. 14.2 for unidirectional motion. The probability of finding an electron at any point along the length is proportional to the square of the amplitude of the wave at that point. This means that, for the free electron with $\lambda = 2L$, the probability of finding it at the midpoint of the length is a maximum. This, however, is not an acceptable result, as the probability of finding the free electron must be the same anywhere within the solid. To overcome this difficulty, the waves are considered to be *travelling waves with a constant velocity*, so that the *time averaged probability of finding an electron is constant* throughout the solid.

In keeping with the minimum energy criterion and the Pauli exclusion principle, the free electrons occupy successive quantum states of increasing kinetic energy. The energy corresponding to the highest filled level at 0 K is called the *Fermi energy* E_F . At 0 K, the free electrons occupy all the levels up to the Fermi level, leaving all those above it empty. At temperatures above 0 K, due to *thermal excitation*, there is a finite probability of some of the electrons from below the Fermi level moving to levels above E_F . This probability is given by the *Fermi-Dirac statistics*, which takes into account the quantum restrictions due to the Pauli exclusion principle. The probability of occupation $P(E)$ of an energy level E by an electron is given by

$$P(E) = \frac{1}{1 + \exp [(E - E_F)/kT]} \quad (14.7)$$

If E_F is independent of temperature, $P(E)$ varies with E as shown in Fig. 14.3. At 0 K, $P(E)$ remains constant at unity with increasing E up to E_F , where it falls abruptly to zero. At $T_1 > 0$ K, some of the electrons just below the Fermi level are thermally excited to higher levels just above E_F . So, $P(E)$ is less than one just below the Fermi level and is greater than zero just above the Fermi level. At a still higher temperature, more electrons leave the lower energy levels and occupy higher levels. The cross-over point, where $P(E)$ is 0.5 for different temperatures in Fig. 14.3, occurs at the same value of the energy level which is E_F . This is true, provided that E_F is independent of temperature, an assumption valid for most ordinary temperatures. Under such conditions, *the Fermi level can be defined as that level which has a 50% probability of occupation by an electron at any temperature.*

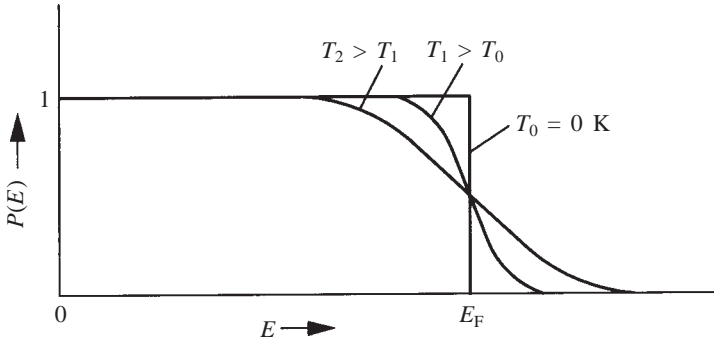


Fig. 14.3 The Fermi–Dirac distribution of free electrons at different temperatures.

14.3 Conduction by Free Electrons

We have already noted that the wave number k takes both positive and negative values. For every electron moving with a certain speed in a direction, there is another electron moving with the same speed in the opposite direction. This equal and opposite velocity distribution in a neutral solid can be biased by an *externally applied electric field* to yield a *net velocity* in one direction. With this biasing, the solid conducts electricity. Under an applied field, the E – k relationship of Fig. 14.1 gets modified to the distribution shown in Fig. 14.4. The negatively charged

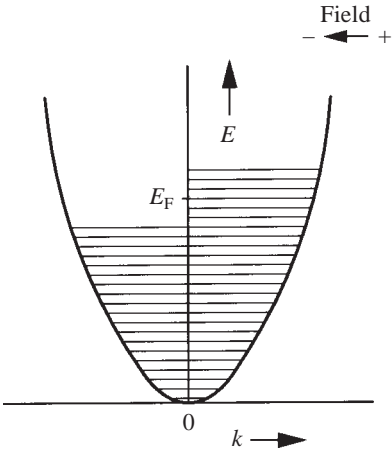


Fig. 14.4 Electrons moving towards the positive end of the applied field acquire extra velocity, while those moving in the opposite direction lose some velocity.

electrons are accelerated towards the positive end of the field. The velocity of the fastest electron moving in the direction of the positive end has a larger magnitude than that of the fastest electron moving towards the negative end of the field. Such redistribution is possible, *only when empty electron states are available immediately above the Fermi level*. This availability is a basic characteristic of conductors, as opposed to semiconductors and insulators.

The force experienced by an electron of charge e in an applied field of gradient \mathcal{E} can be equated to the force as defined in the classical law:

$$\mathcal{E}e = ma \quad (14.8)$$

where m is the mass of the electron and a is the acceleration due to the applied field. The electrons that are accelerated towards the positive end of the field do not continue to increase their velocity indefinitely. They collide with obstacles on their way. Depending on the time interval between two successive collisions, the electrons acquire an average increment of velocity called *drift velocity*, all of which they lose during a collision, as illustrated in Fig. 14.5. The drift velocity is the extra velocity that electrons acquire over and above their normal velocity in the absence of a field.

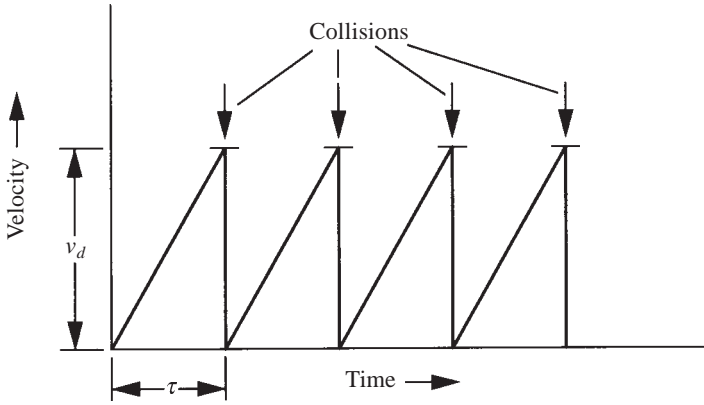


Fig. 14.5 The extra velocity acquired by an electron due to an applied field is lost on collision with an impurity, imperfection or phonon.

If the average collision time is τ and v_d is the drift velocity acquired by the electrons, Eq. (14.8) can be rewritten as

$$m(v_d/\tau) = \mathcal{E}e \quad (14.9)$$

or

$$v_d = \frac{\mathcal{E}e\tau}{m} \quad (14.10)$$

The flux J_e due to the flow of electrons is called the *current density*:

$$J_e = nev_d = \frac{ne^2\tau\mathcal{E}}{m} \quad (14.11)$$

where n is the number of free electrons of charge e . This is in the form of Ohm's law. As conductivity σ is by definition the flux per unit potential gradient, we have

$$\sigma = \frac{ne^2\tau}{m} \quad (14.12)$$

The electrical resistivity ρ is the reciprocal of conductivity.

Example 14.2 Calculate the conductivity of copper at 300 K. The collision time τ for electron scattering is 2×10^{-14} s at this temperature.

Solution The number of free electrons per m^3 of copper is given by

$$n = 6.023 \times 10^{23} \times 8960 / 0.06354 = 8.50 \times 10^{28} \text{ m}^{-3}$$

Electronic charge $e = 1.602 \times 10^{-19}$ C

Mass of a free electron \sim rest mass
 $= 9.1 \times 10^{-31}$ kg

Using Eq. (14.12),

$$\begin{aligned}\sigma &= 8.50 \times 10^{28} \times 1.602^2 \times 10^{-38} \times 2 \times 10^{-14} / (9.1 \times 10^{-31}) \\ &= 4.8 \times 10^7 \text{ ohm}^{-1} \text{ m}^{-1}\end{aligned}$$

The mean free path l of an electron is the mean distance it travels between successive collisions. For an ideal crystal with no impurities and imperfections, the mean free path at 0 K is infinite. That is, there are no collisions and the electrical conductivity is ideally infinite.

Solute atoms provide effective scattering centres for the electrons. Their introduction into the crystal results in collisions, decreasing the mean free path and the conductivity. Similarly, other point imperfections, dislocations and grain boundaries also increase the scattering and decrease the conductivity. Among these scattering centres, impurities *dissolved in the crystal lattice* as solutes are more effective than the others.

At temperatures above 0 K, the atoms vibrate randomly about their mean positions. These vibrations can be considered as elastic waves in the crystal and are called *phonons*. Their random nature destroys the ideal periodicity of a crystal and interferes with the electron motion. Consequently, the mean free path and the conductivity decrease with increasing temperature. At low temperatures, the mean free path and the collision time are proportional to the cube of the reciprocal of temperature in kelvin. At higher temperatures, the vibrational frequency tends to become constant. Here, the mean free path and the collision time are proportional to the reciprocal of temperature.

The electrical resistivity of pure copper with a low density of imperfections is shown as a function of temperature in Fig. 14.6. The variation is cubic at low temperatures and is linear at higher temperatures. If nickel is introduced in the copper lattice as a solute, the resistivity of the alloy is higher at all temperatures than that of pure copper. The increase in resistivity is proportional to the amount of nickel added. The resistivities of Cu-2% Ni and Cu-3% Ni are shown in Fig. 14.6. The resistivities of the Cu-Ni alloys at 0 K have a finite residual value, in contrast to pure copper, where the resistivity is tending to zero at 0 K. This residual resistivity is attributable to the scattering centres provided by the solute atoms which are present even at 0 K, where there is no phonon scattering. The *additivity rule of Mattheissen* describes this behaviour:

$$\rho = \rho_T + \rho_r \quad (14.13)$$

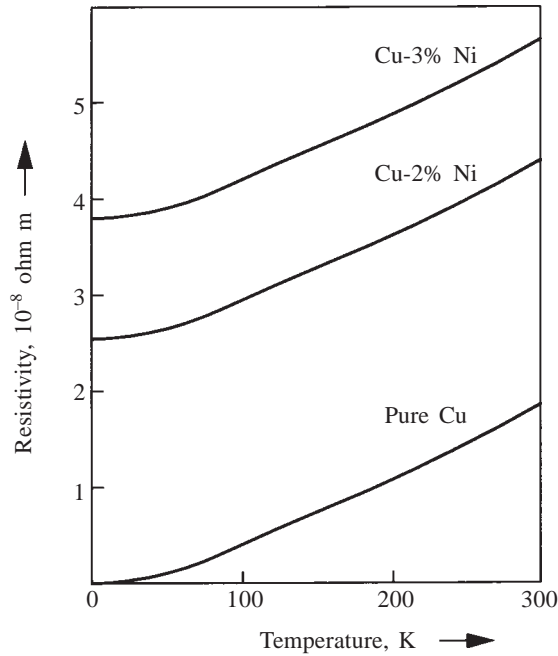


Fig. 14.6 The electrical resistivity of pure copper and Cu–Ni alloys as a function of temperature.

where ρ_T is the thermal part of the resistivity and ρ_r is the residual resistivity due to solute atoms and other imperfections in the crystal.

Example 14.3 The resistivity of pure copper at room temperature is 1.8×10^{-8} ohm m. The resistivity of Cu-4% Ni alloy at room temperature is 7.0×10^{-8} ohm m. Estimate the resistivity due to impurity scattering by 1% of nickel in the copper lattice.

Solution Taking the resistivity of copper at 0 K to be negligible, 1.8×10^{-8} ohm m is the thermal part of resistivity of copper as well as of the alloy between 0 K and room temperature. The impurity scattering by 1% of nickel is then equal to $(7.0 - 1.8) \times 10^{-8}/4 = 1.3 \times 10^{-8}$ ohm m.

14.4 Conductor and Resistor Materials

The resistivity, the temperature coefficient of resistance, the density and the tensile strength of typical conductor and resistor materials are listed in Table 14.2.

TABLE 14.2
Properties of Typical Conductors and Resistors at Room Temperature

Material	Resistivity, 10^{-8} ohm m	Temperature coefficient α , K^{-1}	Density, 10^3 kg m $^{-3}$	Tensile strength*, MN m $^{-2}$
Silver	1.5	0.0040	10.49	125
Copper	1.7	0.0043	8.96	210
Gold	2.2	0.0035	19.32	138
Aluminium	2.8	0.0042	2.70	60
Tungsten wire	5.5	0.0045	19.3	2800
Molybdenum wire	4.9	0.0050	10.2	700
Platinum wire	10.9	0.0037	21.45	350
Tantalum wire	15.5	0.0032	16.6	490
Nichrome wire	108	0.0001	8.41	1000
Manganin	48	0.00002	8.2	420
Kanthal wire	135	0.00003	7.2	800

*The tensile strength values given here are approximate, as they depend on the prior thermal and mechanical history of the metal.

For use as *conductors* in applications such as transmission lines and distribution lines, low I^2R loss is the primary consideration and the choice would be from amongst the best conductors, keeping in view the cost, fabricability and mechanical strength. *Copper and aluminium* are the most likely choices. For long distance transmission lines, aluminium is chosen. As a large cross section would reduce the I^2R loss, thick cables are preferred. If the elastic modulus of the aluminium cables is improved by reinforcement with steel as in ACSR (aluminium conductor steel reinforced) cables, the distance between successive poles along the transmission line can be substantially increased. More expensive copper is used for distribution lines, busbars and other energy conversion applications. OFHC (oxygen-free high conductivity) copper is often specified. Among the common solutes in copper, Fe, P and As are the most harmful in impairing the electrical conductivity.

For *electrical contacts* used in switches, brushes and relays, the material must possess high electrical conductivity, high thermal conductivity, high melting point and good oxidation resistance. High thermal conductivity helps to dissipate the heat effectively. High melting point is desirable so that any accidental overheating does not fuse together the contact points. Good oxidation resistance is necessary to keep the contact clean and free of insulating oxides. Copper and silver largely satisfy the above requirements. For low cost, copper is commonly used. For critical contacts such as those used in aircrafts, silver is preferred. The low mechanical strength of pure silver is increased by the dispersion of fine particles of CdO. Dislocations moving in silver have to bend

around the dispersed CdO particles and a fine dispersion increases the strength. CdO improves the wear resistance of silver. It also decomposes at the melting point of silver, thereby absorbing much of the heat generated by arcing and minimizing the loss of expensive silver by evaporation.

For *resistor* applications, the primary requirements are uniform resistivity (achieved in a homogeneous alloy), stable resistance (achieved by avoiding metallurgical changes such as ageing and relaxation of residual stresses), small temperature coefficient of resistance (α) and low thermoelectric potential with respect to copper. A small α minimizes the error in measurements due to variations in the ambient temperature. α is defined as

$$\alpha = \frac{1}{R} \frac{dR}{dT} \quad (14.14)$$

where R is the resistance of the alloy at temperature T . For manganin alloy (87% Cu and 13% Mn), α is only $20 \times 10^{-6} \text{ K}^{-1}$ as against $4000 \times 10^{-6} \text{ K}^{-1}$ for pure copper. Constantan (60% Cu and 40% Ni) is another such alloy. These alloys have also good resistance to atmospheric corrosion, another desirable property in a resistor.

A low thermoelectric potential with respect to copper, to which the resistor is commonly connected, reduces errors due to temperature differences between junctions. For high precision, dissimilar junctions should be maintained at the same temperature so that no thermoelectric potential develops.

Ballast resistors are used to maintain constant current in some industrial circuits. If the flow of current increases, the temperature increases and the resistance of the ballast increases. This in turn decreases the current in the circuit towards the initial value. An iron–nickel alloy with 71% Fe and 29% Ni with excellent oxidation resistance and a high α of $4500 \times 10^{-6} \text{ K}^{-1}$ is used for this application.

For *heating elements*, the primary requirements are high melting point, high electrical resistance, good oxidation resistance, good creep strength, low elastic modulus and low thermal expansion. The last two requirements help in reducing thermal fatigue due to repeated heating and cooling. The heating elements should be designed in a way as to allow unhindered expansion and contraction, for example, in the form of a coil of wire. Nichrome (80% Ni and 20% Cr) and Kanthal (69% Fe, 23% Cr, 6% Al and 2% Co) are used for heating elements up to 1300°C . SiC and MoSi_2 can be used at higher temperatures up to 1700°C . Graphite, by virtue of its high sublimation temperature and good fabrication properties, is also widely used up to 1800°C . Molybdenum and tantalum need protective atmospheres at high temperatures, as their oxidation resistance is poor. By virtue of its very high melting point (3410°C), tungsten is used for filaments in incandescent lamps. Its creep resistance at white heat (above 1500°C) can be improved by dispersion hardening with thorium (ThO_2).

Resistance thermometers should have a high temperature coefficient of resistance for good sensitivity. A pure metal is obviously the choice for this application. Platinum, obtainable in very pure form, is used.

14.5 Superconducting Materials

14.5.1 The Superconducting Phenomenon

Referring to Fig. 14.7, the resistivity of pure silver decreases gradually to a low but measurable value of $\sim 10^{-11}$ ohm m at 0 K. Silver is not a superconductor.

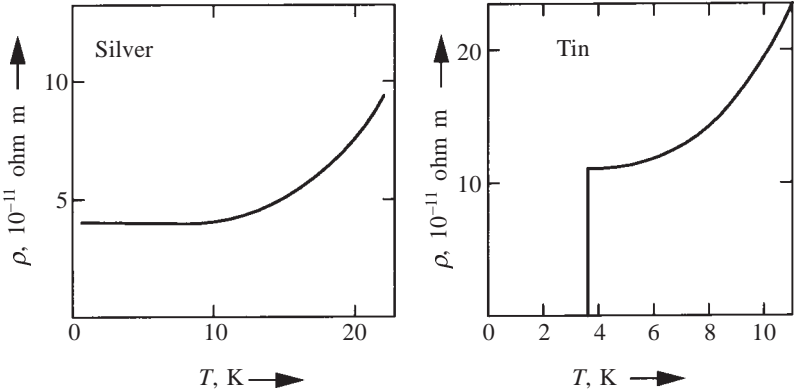


Fig. 14.7 The electrical resistivity ρ of (a) pure silver, and (b) tin, as a function of temperature near 0 K.

The resistivity of tin, on the other hand, falls abruptly at about 4 K to a value below the measurable limit of 10^{-20} ohm m. The temperature at which this abrupt fall occurs is called the *superconducting transition temperature* T_c . Tin is in the normal state above T_c and in the superconducting state below T_c . The resistivity in the superconducting state is zero for all practical purposes. A current induced in a superconducting ring persists for years with negligible decay.

The first application that was conceived of the superconducting effect was to produce a large permanent magnetic field by inducing a perpetual current in a superconductor. However, at temperatures below T_c , as the magnetic field strength reaches a critical value H_c , the superconductivity disappears. This is illustrated in Fig. 14.8. At $T = T_c$, $H_c = 0$. H_c increases as the temperature decreases below T_c . A similar graph relates the critical current density J_c that a superconductor can carry at temperatures below T_c . At $T = T_c$, $J_c = 0$. J_c increases as the temperature decreases below T_c in a manner similar to H_c in Fig. 14.8. In other words, the maximum current that a superconductor carries at a given temperature below T_c is limited by the magnetic field it produces at the surface of the superconductor.

In the superconducting state, the flux lines of a magnetic field are ejected out of the superconductor, as illustrated in Fig. 14.9. This effect is known as the *Meissner effect*. A superconductor is a perfect diamagnetic material, with the magnetic susceptibility χ equal to -1 .

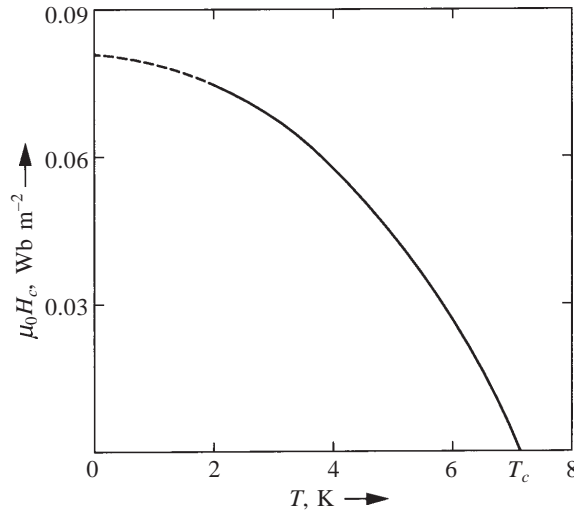


Fig. 14.8 The critical magnetic field H_c as a function of temperature for the superconducting-to-normal transition. $H_c = 0$ at T_c .

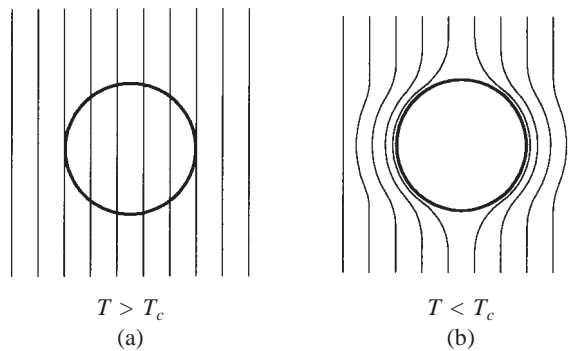


Fig. 14.9 The flux lines of an applied magnetic field are ejected out of the solid on crossing to the superconducting state.

The superconducting effect has been explained in the Bardeen–Cooper–Schreiffer (BCS) theory as a three-way interaction between two electrons and a phonon (a phonon is the quantum of energy in an elastic wave, analogous to a photon, which is the quantum of energy in an electromagnetic wave). The phonon scattering due to lattice vibrations felt by electron 1 is exactly nullified by electron 2, such that the electron pair moves through the lattice without feeling any scattering effect from lattice vibrations. This attractive interaction energy is more than the repulsive energy between the two electrons; yet it is small enough in magnitude to be disrupted by the thermal energy available at $T > T_c$.

14.5.2 **Type I and Type II Superconductors**

Type I or the ideal superconductor when placed in a magnetic field repels the flux lines totally, till the magnetic field attains the critical value H_c . The magnetization M is equal to $-H$ up to T_c , where it drops to zero, as shown in Fig. 14.10a. Type II or hard superconductors are those in which the ideal behaviour is seen up to a lower critical field H_{c1} , beyond which the magnetization gradually changes and attains zero at an upper critical field designated H_{c2} , see Fig. 14.10b. The Meissner effect is incomplete in the region between H_{c1} and H_{c2} ; this region is known as the *vortex region*. The normal behaviour is observed only beyond H_{c2} . The magnetic flux lines gradually penetrate the solid, as the field is increased beyond H_{c1} and the penetration is complete only at H_{c2} .

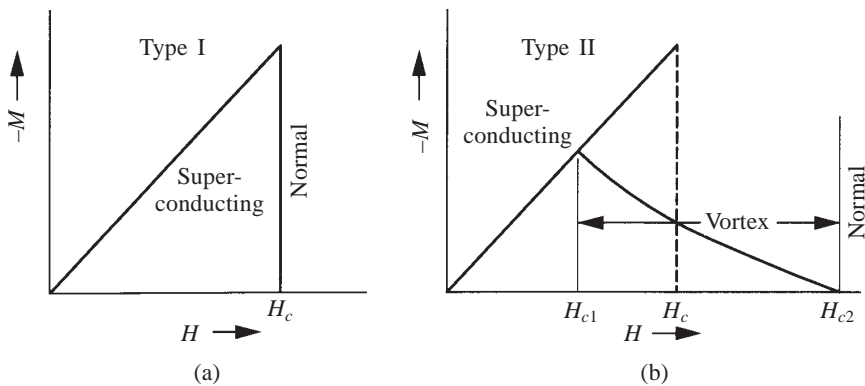


Fig. 14.10 The magnetization M versus the critical magnetic field H_c for (a) Type I and (b) Type II superconductors.

Type II superconductors are of great practical interest because of the high current densities that they can carry. The type II state is determined by the microstructural condition of the material. Heavily cold-worked and recovery annealed material have cell walls of high dislocation density and this microstructure effectively pins the magnetic flux lines and makes their penetration difficult. Grain boundaries also exert a pinning action; an extremely fine grain size is effective in increasing H_{c2} . Similarly, the dispersion of very fine precipitates in the matrix with the interparticle spacing of about 300 Å results in optimum flux pinning. The critical current density J_c also increases, as H_{c2} increases. The very high current densities obtainable in a Nb-40% Ti alloy at 4.2 K at a magnetic field strength of $0.9H_{c2}$ are listed below as a function of the microstructural condition.

Microstructure	$J_c, \text{ A m}^{-2}$
Recrystallized	10^5
Cold worked and recovery annealed	10^7
Cold worked and precipitation hardened	10^8

14.5.3 Potential Applications

Superconducting materials are already in use for producing very strong magnetic fields of about 50 Tesla, which is much larger than the field obtainable from an electromagnet. Such high magnetic fields are required in MHD power generators. At high magnetic field strengths, a conventional copper solenoid consumes about 3 MW, whereas a superconducting magnet consumes about 10 kW. In addition, the copper solenoid will require about 2000 l/min of water circulation to avoid burning down due to Joule heating. Magnetic energy can be stored in large superconductors and drawn as required, to counter voltage fluctuations during peak loading.

Superconductors can be used to perform logic and storage functions in computers. A Josephson junction consists of a thin layer of insulating material between two superconducting solids. The unique current-voltage characteristics associated with the Josephson junction are suitable for memory elements. Switching times of the order of 10 ps (10^{-11} s) have been measured.

Arising from the Meissner effect, a superconducting material can be suspended in air against the repulsive force from a permanent magnet. This magnetic levitation effect can be used in transportation. As there is no I^2R loss in a superconductor, power can be transmitted through superconducting cables without loss. With the liquid N_2 environment, the economics of such transmission has become more favourable for adoption.

14.5.4 New Developments

Till the year 1986, the highest known transition temperature T_c was 23 K in the Nb_3Ge alloy. In 1986, Bednorz and Mueller reported a significant increase in T_c to 34 K in a La-Ba-Cu-O ceramic material. This Nobel prize-winning discovery was soon followed by further big increases in the transition temperature. The oxide with the nominal formula $YBa_2Cu_3O_{7-x}$ has a transition temperature of ~ 90 K.* This transition temperature is 13 K above the boiling point of liquid N_2 (77 K). Compare this with the boiling point of liquid He (4 K) and that of liquid H_2 (23 K), which is a safety hazard. That a superconductor can function in liquid nitrogen is itself a remarkable achievement. A liquid nitrogen environment is far easier and cheaper to obtain than a liquid helium medium. Further, it takes about 25 times more energy to cool from 77 K to 4 K than from room temperature to 77 K.

The oxide $YBa_2Cu_3O_{7-x}$ is prepared by heating compacted powder mixtures of Y_2O_3 , $BaCO_3$ and CuO in the right proportion to temperatures between 900 and 1100°C. $BaCO_3$ decomposes at this temperature to BaO and CO_2 . This is often followed by another annealing treatment at 800°C in an atmosphere of oxygen. The heat treatment conditions such as the partial pressure of oxygen in the atmosphere are critical for obtaining a high T_c . The crystal structure of the powder product obtained is related to the cubic perovskite structure as illustrated in Fig. 14.11. Three body-centered cubic unit cells are stacked one above another. The atom distributions in the unit cells are as follows:

*New compounds such as $Tl(Bi)-Ba(Sr)-Ca-Cu-O$ are now known with a reproducible T_c of about 125 K.

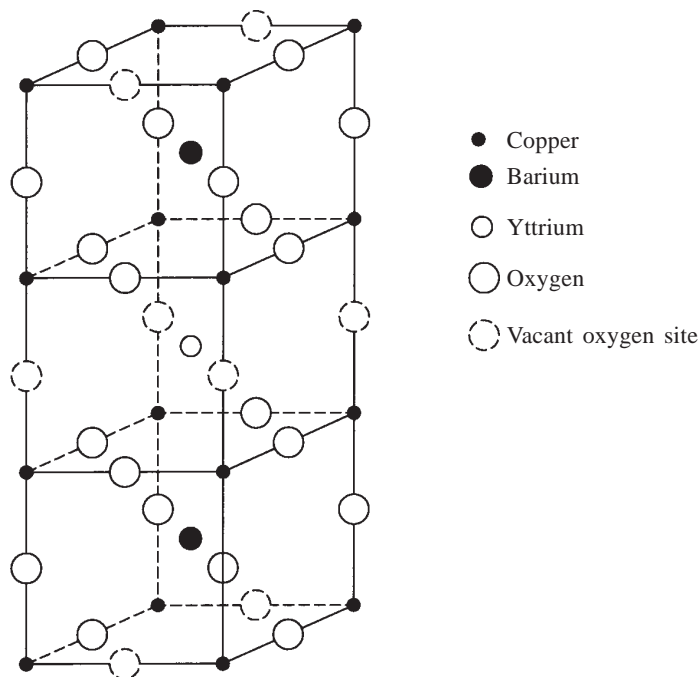


Fig. 14.11 The crystal structure of yttrium–barium–copper oxide.

Cu at the body corners : $8 \times 1/8 \times 3 = 3$

Ba at two body centres : $1 \times 2 = 2$

Y at one body centre : $1 \times 1 = 1$

O at mid-points of edges : $12 \times 1/4 \times 3 = 9$

Some oxygen positions are vacant and their number is between 6.5 and 7. The superconducting properties appear to be a sensitive function of the oxygen content and, therefore, of the partial pressure of oxygen during heat treatment.

The engineering aspects of this oxide superconductor still remain elusive. It is reactive, brittle, unable to support any significant stress and cannot be easily formed or joined. Also, the superconducting properties deteriorate during heating for forming purposes or even in a humid room. Explosive forming and isotactic pressing are two forming processes that appear to be promising. In explosive forming, the oxide powder is placed between two copper plates in a container, which is filled with an explosive. On detonating, shock waves are generated and compressive pressures up to 50 000 atm are exerted on the composite sandwich-like arrangement, which gets compacted into one solid mass. The temperature attained in the chamber is less than 100°C. The very high compressive stress plastically deforms the oxide, which should help to increase its critical-current carrying capacity. Once encapsulated, the outer copper layer provides for easy external connection and also ensures that there is no deterioration of the superconducting properties due to environmental effects.

SUMMARY

1. Electrical resistivity of materials covers a very wide range, over 25 orders of magnitude.
2. In a solid where the potential field is uniform, the total energy of the free electrons can be taken to be their kinetic energy. This energy varies as the square of the wave number k .
3. The Fermi energy level is defined as that level which has a 50% probability of occupation by an electron. At 0 K, the Fermi level is the highest filled electron energy level.
4. At temperatures above 0 K, due to thermal excitation, the distribution of electrons near the Fermi level changes as given by the Fermi–Dirac statistics.
5. The availability of empty electron states immediately above the Fermi level is a basic requirement for electrical conduction.
6. Ohm's law arises, as conducting electrons acquire a drift velocity under an applied field, due to collisions with impurities, imperfections and phonons.
7. For high conductivity and large temperature coefficient of resistance, α , pure metals are chosen. For heating purposes by I^2R dissipation and for low temperature coefficient of resistance, alloys are more suitable.
8. The transition temperature T_c , the critical magnetic fields H_{c1} and H_{c2} , and the critical current density J_c are the main properties of interest of superconductors. Among these, J_c is structure sensitive and can be maximized by proper control of the microstructure.
9. The recent discovery of high T_c oxide superconductors is an important advance in scientific research, with far reaching technological implications.

PROBLEMS

- 14.1** Calculate the kinetic energy and the de Broglie wavelength of an electron that falls through a potential of (i) 500, (ii) 5000, and (iii) 50000 V.
Answer: (i) 500 eV, 0.55 Å, (ii) 5000 eV, 0.17 Å, and (iii) 50000 eV, 0.055 Å
- 14.2** Derive the kinetic energy of free electrons as a function of their wave number.
- 14.3** What should be the energy of the quantum state, which has 19 for the sum of the squares of the quantum numbers n_x , n_y and n_z ? What is the degeneracy of this state? Assume the solid to be 10 mm \times 10 mm \times 10 mm.
Answer: 1.15×10^{-32} J; 24.

- 14.4** Show that the probability of occupancy of energy level E by an electron is (i) negligible for $E \gg E_F$, (ii) 0.5 for $E = E_F$, and (iii) ~ 1 for $E \ll E_F$.
- 14.5** The Fermi level for potassium is 2.1 eV. Calculate the velocity of the electrons at the Fermi level.
Answer: $8.6 \times 10^5 \text{ m s}^{-1}$.
- 14.6** The Fermi level of silver is 5.5 eV. Calculate the fraction of free electrons at room temperature located up to a width of kT on either side of E_F .
Answer: 0.01.
- 14.7** The resistance of a sample of copper wire of 0.1 mm diameter and 0.2 m long is 0.439 ohm. Calculate its resistivity and conductivity.
Answer: $1.72 \times 10^{-8} \text{ ohm m}$ and $5.81 \times 10^7 \text{ ohm}^{-1} \text{ m}^{-1}$.
- 14.8** The resistivity of silver at room temperature is $1.6 \times 10^{-8} \text{ ohm m}$. Calculate the collision time for electron scattering.
Answer: $3.8 \times 10^{-14} \text{ s}$.
- 14.9** The resistivity of aluminium at room temperature is $2.62 \times 10^{-8} \text{ ohm m}$. Calculate the incremental velocity acquired by the free electrons in a field gradient of 100 V m^{-1} .
Answer: 0.39 m s^{-1} .
- 14.10** Calculate the per cent increase in resistivity of (i) copper and (ii) nichrome on heating from 300 K to 1000 K.
Answer: (i) 300%, (ii) 7%.
- 14.11** If the ambient temperature varies by $\pm 10^\circ\text{C}$ about the mean value of 20°C , what is the maximum possible error between two measurements of resistance made at different times of (i) a copper resistor, and (ii) a manganin resistor?
Answer: (i) 8.6%, (ii) 0.04%
- 14.12** Resistivity changes of 1 in 10^4 can be detected by means of a bridge. What is the smallest change in temperature that you can measure with a platinum thermometer at -200°C . At this temperature, $\alpha = 0.03 \text{ K}^{-1}$.
Answer: 0.0033°C .
- 14.13** Explain why nichrome and not copper is used as a heating element.
- 14.14** A wire whose diameter is 2 mm must carry a 20 A current in an application. The maximum power dissipation along the wire is 5 W m^{-1} . Calculate the minimum permissible conductivity of the wire for this application.
Answer: $2.55 \times 10^7 \text{ ohm}^{-1} \text{ m}^{-1}$.
- 14.15** Estimate the mean free path of free electrons in pure copper at 4 K. The collision time for phonon scattering at 4 K is 10^{-9} s . The Fermi energy level for copper is 7 eV.
Answer: 1.57 mm.

- 14.16** Explain why aluminium used in long distance transmission lines cannot be strengthened by solid solution.

MULTIPLE CHOICE QUESTIONS

- The SI units of electrical conductivity are
A. $\text{kg}^{-1} \text{m}^{-3} \text{s}^3 \text{A}^2$ B. $\text{kg} \text{m}^3 \text{s}^{-3} \text{A}^{-2}$ C. ohm m D. $\text{ohm}^{-1} \text{m}^{-3}$
- The correct order of increasing resistivity of the following is
A. nickel, doped silicon, sodium silicate, pure silica
B. doped silicon, pure silica, nickel, sodium silicate
C. pure silica, doped silicon, sodium silicate, nickel
D. nickel, pure silica, doped silicon, sodium silicate
- Volt in SI units is
A. W A B. $\text{N m}^2 \text{s}^{-1} \text{A}^{-1}$ C. $\text{kg m}^2 \text{s}^{-3} \text{A}^{-1}$ D. ohm A^{-1}
- The degeneracy of the quantum states with $(n_x^2 + n_y^2 + n_z^2) = 6$ is
A. 8 B. 12 C. 24 D. 48
- The energy level difference between two successive levels for the lowest energy free electrons is in order of magnitude
A. 1 eV B. 10^{-33} J C. 10^{-23} J D. 10^{-19} J
- In a metal of $1 \text{ cm} \times 1 \text{ cm} \times 1 \text{ cm}$, the lowest energy free electron has a wavelength of
A. 0.5 cm B. 1 cm
C. 2 cm D. need Fermi energy to compute
- The probability of finding a free electron in Ag at 300 K at an energy level $1.01E_F$ is given by (E_F for Ag = 5.5 eV)
A. 0.5 B. 0.11 C. 0.005 D. 0
- If the Fermi energy of silver is 5.5 eV, the wave number of the fastest electron at 0 K has the magnitude (in m^{-1})
A. 0.85×10^{10} B. 7.54×10^{10} C. 1.20×10^{10} D. 0.19×10^{10}
- The Fermi level for Cu is 7 eV. The maximum velocity of free electrons at 0 K is
A. 1570 km s^{-1} B. 1110 km s^{-1} C. 860 km s^{-1} D. 0 km s^{-1}
- The acceleration (m s^{-2}) of a free electron in an electric field of 100 V m^{-1} is
A. 1.76×10^3 B. 1.1×10^{13} C. 1.76×10^{13} D. 1.1×10^{32}

23. High conductivity aluminium should not have
 - A. steel rod reinforcement
 - B. solute atoms such as Cu, Ag and Au
 - C. high dislocation density
 - D. dissolved impurities
24. The electrical resistivity of very pure silver near 0 K is
 - A. not measurable
 - B. below 10^{-20} ohm m
 - C. very low but measurable
 - D. very high
25. Tick the materials with the approximately correct superconducting transition temperature given in brackets
 - A. Sn (4 K)
 - B. Nb₃Ge (23 K)
 - C. Y–Ba–Cu oxide (90 K)
 - D. Y–Ba–Cu oxide (300 K)
26. The magnetization of a superconductor is
 - A. 0
 - B. $-B$
 - C. -1
 - D. $-H$
27. The atoms that are located at the body centres of the cubic unit cells of the YBa₂Cu₃O_{7-x} superconductor are
 - A. Y
 - B. Ba
 - C. Cu
 - D. O
28. The first measured T_c in a ceramic superconductor by Bednorz and Mueller was
 - A. 4 K
 - B. 23 K
 - C. 34 K
 - D. 90 K
29. The following microstructural features can improve J_c in a superconductor
 - A. dislocation tangles
 - B. grain boundaries
 - C. fine precipitates
 - D. low test temperature
30. Explosive forming of a ceramic superconductor causes
 - A. J_c to increase
 - B. T_c to increase
 - C. plastic deformation
 - D. brittle fracture of the ceramic
31. Cold working of oxide superconductors is possible with
 - A. high hydrostatic pressure
 - B. moderate tensile stress
 - C. large tensile force
 - D. none of these
32. Switching times with a Josephson junction are in order of magnitude
 - A. 10^{-2} ns
 - B. 0.1 μ s
 - C. 10^{-15} s
 - D. 1 μ s
33. The room temperature electrical resistivity (ohm-m) of the new oxide superconductors lies around
 - A. 10^{-9}
 - B. 10^{-5}
 - C. 10^7
 - D. 10^{18}
34. The critical current density J_c that a superconductor can carry depends on
 - A. T with respect to T_c
 - B. dislocation density
 - C. precipitate distribution
 - D. none of these

Answers

- | | | | | |
|-------|----------|-------------|-------------|-------------|
| 1. A | 2. A | 3. C | 4. C | 5. B |
| 6. C | 7. B | 8. C | 9. A | 10. C |
| 11. D | 12. A, B | 13. C | 14. C | 15. B |
| 16. A | 17. C | 18. C | 19. C | 20. C |
| 21. D | 22. D | 23. B, C, D | 24. C | 25. A, B, C |
| 26. D | 27. A, B | 28. C | 29. A, B, C | 30. A, C |
| 31. A | 32. A | 33. B | 34. A, B, C | |

Sources for Experimental Data

G.W.A. Dummer, *Materials for Conductive and Resistive Functions*, Hayden Book Co., New York (1971).

Metals Handbook, 10th ed., Vol. 3, *Special Purpose Materials*, ASM International, Materials Park, Ohio (1990), pp. 804–1024.

Suggestions for Further Reading

C. Kittel, *Introduction to Solid State Physics*, Wiley, New York (1976), Chaps. 6 and 12.

L.E. Murr, A.W. Hare and N.G. Eror, Introducing the Metal Matrix High Temperature Superconductor, in *Advanced Materials and Processes*, ASM International, Materials Park, Ohio (1987), Vol. 145(10), pp. 36–44.

CHAPTER

15 Semiconductors

A number of solid state devices have come into use in the last several decades, using semiconductor crystals. The field of electronics has been revolutionized since the discovery of the transistor in 1948. The fragile vacuum tubes have now been largely replaced by solid state diodes and triodes. This technological progress has been in a large measure due to improved techniques in the preparation, purification and characterization of the raw materials needed to make the solid state devices.

In this chapter, we shall consider briefly the energy band model for solids in general and for pure and doped semiconductors in particular. We shall also deal with common semiconductor materials and their fabrication. The last section is devoted to a description of some common semiconductor devices.

Units

Quantity	SI units		Other units
	<i>Unit</i>	<i>Symbol</i>	
de Broglie wavelength λ	metre	m	Å
Wave number k	per metre	m ⁻¹	–
Energy gap E_g	joule	J	eV
Contact potential eV_0	joule	J	eV
Concentration of conduction electrons or holes, n_e or n_h	per cubic metre	m ⁻³	–
Mobility μ_e or μ_h	metre squared per volt per second	m ² V ⁻¹ s ⁻¹	–
Conductivity σ	per ohm per metre	ohm ⁻¹ m ⁻¹	–

Constants

Boltzmann's constant k	$= 1.380 \times 10^{-23} \text{ J K}^{-1}$
Electronic charge e	$= 1.602 \times 10^{-19} \text{ C}$
Electron rest mass m_0	$= 9.109 \times 10^{-31} \text{ kg}$

15.1 The Energy Gap in Solids

In the model of free electrons discussed in the last chapter, the potential field in which the electrons are moving within the crystal is assumed to be constant. In reality, of course, this is a gross approximation. The potential energy of an electron will be undoubtedly a function of its position with respect to the ion cores and the other electrons in the system. For example, its potential energy will be less if it is close to a (positive) ion core and more than the average when it is farthest away from neighbouring ion cores. This fact combined with the *diffracting properties* of electrons has an important effect on the electron energy calculations, especially at certain critical values of the de Broglie wavelength λ of the electrons.

Consider the motion of electrons along only one axis as before. Electrons moving to the right or to the left are scattered by the atoms in the path. This scattering is largely *coherent* and is being reinforced in the direction of motion of the electrons. If, however, the spacing d between neighbouring atoms along the axis is such that the Bragg equation

$$n\lambda = 2d \sin \theta \quad (15.1)$$

is exactly satisfied, the electrons are totally reflected. For the unidirectional motion under consideration, the atoms along the axis can be visualized as successive 'planes' perpendicular to the direction of electron motion so that $\sin \theta = 1$. Then, from Eq. (15.1), the wavelength λ of the electron takes on critical values of $2d$, d , $2d/3$, and so on, for successively increasing order of reflection, denoted by n . Expressed in terms of the wave number k ($k = 2\pi/\lambda$), we get the critical conditions at $k = \pm\pi/d, \pm 2\pi/d, \pm 3\pi/d, \dots$. The electrons are reflected back and forth repeatedly, when they have the critical wave number. We can say that the net velocity of the electrons in a given direction is zero, due to repeated reflection.

The wave form of the electrons under such critical conditions comes out to be *two standing waves* rather than the travelling wave we visualized for the free electron. As pointed out in the last chapter, the travelling wave gives a constant, time-averaged density of electrons throughout the solid. The standing wave, on the other hand, gives a *periodic variation in the amplitude* and in the electron probability density in the crystal. Standing waves can be obtained from linear combinations of two travelling waves moving in opposite directions, consistent with the repeated change in the direction of electron motion due to reflection. Two possible combinations (the sum and the difference) of the travelling waves give respectively a cosine wave and a sine wave, see Fig. 15.1. These waves have the maximum amplitude at fixed points in space.

The probability of finding an electron at a point is proportional to the square of the amplitude of the wave at that point. The probability density ρ corresponding to the cosine and the sine waves is shown in the middle in Fig. 15.1. (For comparison, the constant, time-averaged probability of the travelling wave is also shown.) The electron is most likely to be found at the crest of the probability density curve. In other words, the electron is *localized* here, in contrast to the free electron, which can be found anywhere within the crystal.

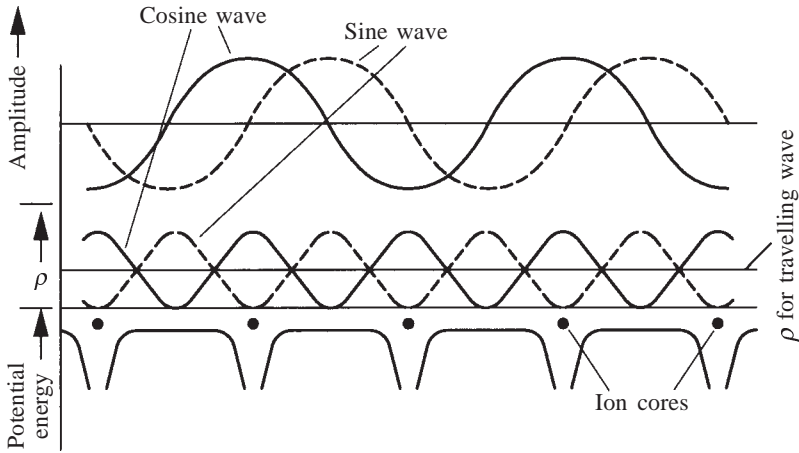


Fig. 15.1 The wave form of electrons are two standing waves when the Bragg law is satisfied. The electron probability density ρ is shown for the two waves. At the bottom, the potential energy variation as a function of electron position is sketched.

As pointed out earlier, the potential energy of an electron is a function of its position in the crystal. When the electrons are localized, the potential energy variations due to location have also to be taken into account. We cannot assume any longer the potential energy to be zero and use the kinetic energy in place of the total energy, as we did in the case of the free electrons. The E - k relationship shown in Fig. 14.1 is replotted in Fig. 15.2, taking into account the deviations from the average potential energy. When there is no critical condition and the electrons are not reflected, they have the average potential energy, which can be taken to be zero. At a critical condition, however, the electron is described by a

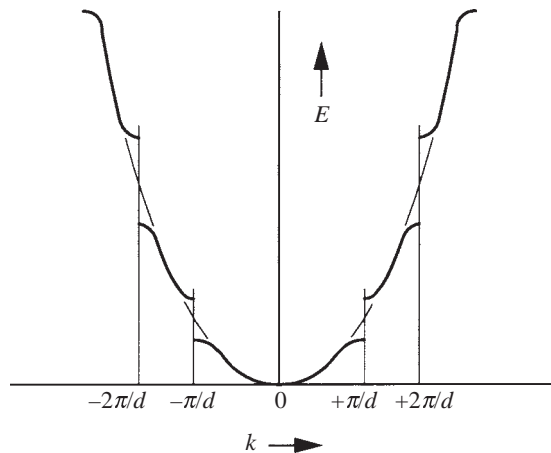


Fig. 15.2 E - k plots show a break in the curve at the critical values of the wave number k .

standing wave and is therefore localized. When the wave form is a standing cosine wave, the electron is close to an ion core (Fig. 15.1). Therefore, it has a lower potential energy than the average. When the wave form is a standing sine wave, the electron is farthest away from the ion cores and correspondingly has a potential energy greater than the average. These deviations from the average potential energy appear as a break in the E - k curve. The break occurring at a critical value of k gives rise to an *energy gap*. Two groups of very closely spaced energy levels called *energy bands* are separated by an energy gap at the critical value of k . The magnitude of this energy gap is the difference in the potential energy of the two electron locations.

When there is no critical condition, E varies as k^2 . So, the slope of the E - k curve increases linearly with k . As the critical value of k is approached, the slope decreases and becomes zero at the critical value. Starting from the inflection point on the curve, the diffraction effects become increasingly important, slowing down the electron continuously to a net velocity of zero at the critical value. This is a strange effect in the sense that with increasing kinetic energy of the electron, its velocity is decreasing! The effective mass of the electron, denoted by m^* , is negative in the region between the inflection point and the critical condition in the E - k curve.

In the three-dimensional crystal, the electrons can move in any direction not necessarily normal to the reflecting planes. The Bragg angle θ can have any value from 0 to 90°. Critical values of k will then correspond to $\pm n\pi/(d \sin \theta)$ and not $\pm n\pi/d$, which we used for the one-dimensional case. Depending on the value of θ , reflection will occur at different values of k , even for the same set of reflecting planes. Consider, for example, the first order reflection from a set of parallel (100) planes in a cubic crystal. The critical condition for this case is $k = \pm\pi/(d \sin \theta)$, where d is the spacing between successive (100) planes. An electron moving in a [100] direction, which is perpendicular to the (100) planes, will have the critical condition at $k = \pi/d$, as $\sin \theta$ is one here. If the direction of electron motion is gradually changed towards [110], the Bragg angle for reflection from (100) planes decreases from 90°. Correspondingly, k increases to larger values than the minimum of π/d . When the electron motion is along [110], the Bragg angle is 45° and $k = \pi/(d \sin 45^\circ) = \sqrt{2}\pi/d$. If the electron direction is further changed towards [010], k does not increase further, as reflection now occurs from (010) planes instead of (100) planes. As (010) is in the same family as (100), d is the same for both. It is easy to see that the critical condition for (010) reflection is $k = \sqrt{2}\pi/d$, when the direction of electron motion is [110] and decreases to π/d , as the direction of motion is changed to [010]. Figure 15.3 illustrates this point. For all directions of electron motion, $\sqrt{2}\pi/d$ is the maximum value of k for first order reflection from the family of {100} planes.

The E - k curves for electron motion along [100] and [110] directions are shown in Fig. 15.4 for first order reflection from (100) planes. The gap in the E - k curve occurs at $k = \pi/d$ for [100] motion and at $k = \sqrt{2}\pi/d$ for [110] motion. The *effective gap* is between the bottom line of the gap in the [110]

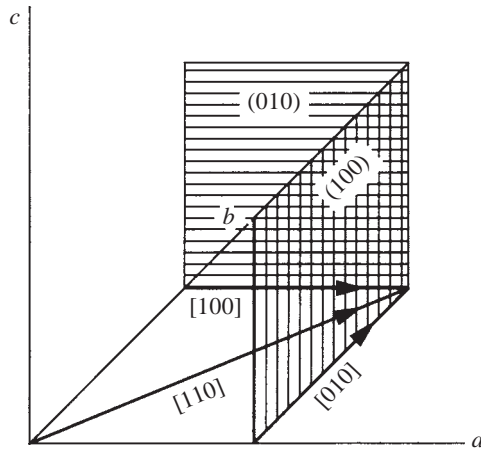


Fig. 15.3 Direction of electron motion and reflecting planes in a cubic crystal.

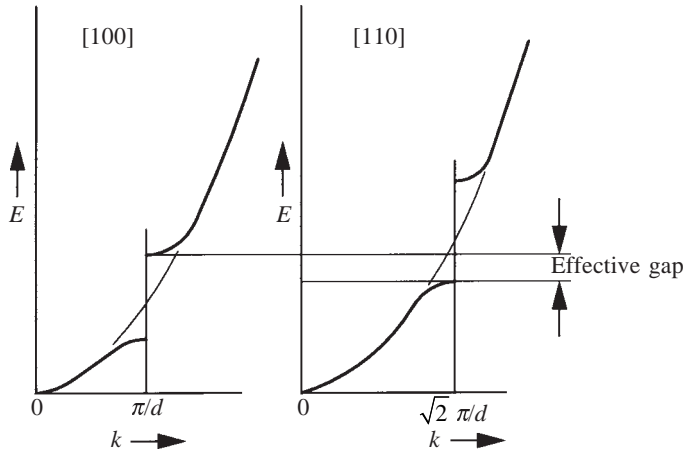


Fig. 15.4 The effective forbidden gap refers to those energy values which an electron cannot take, whatever be the direction of motion in the crystal.

direction and the top line of the gap in the $[100]$ direction, as illustrated in Fig. 15.4. This effective gap is called the *forbidden gap*, because the electrons cannot take those values of energy that lie in the forbidden gap, regardless of their direction of motion. The forbidden gap disappears, if the bottom line of the gap for $[110]$ direction lies above the top line of the gap in the $[100]$ direction. Such a situation will arise, when the potential energy of an electron is not a strong function of location in a crystal. The energy bands on either side of a critical value of k are said to overlap and there is no forbidden energy.

Figure 15.2 is the E - k plot for unidirectional motion of electrons. When all directions of electron motion and all possible reflections are considered together, the ranges of k that lie between the forbidden gaps are called *Brillouin zones*. The first

Brillouin zone contains all the energy levels up to the first E - k discontinuity, the second zone contains all levels between the first and the second discontinuities and so on. The first zone for an FCC crystal is a polyhedra bounded by planar surfaces of the $\{111\}$ and $\{200\}$ types, these two corresponding to the first and the second allowed reflections for the FCC lattice.

The simple energy band representation is adequate for our further discussion. The outermost energy band that is full or partially filled is called the *valence band* in solids. The band that is above the valence band and that is empty at 0 K is called the *conduction band*. Solids are classified on the basis of their band structure as metals, semiconductors and insulators.

Metals are those solids which have vacant electron energy states immediately above the highest filled level of the valence band. This can happen in two ways. In the first case, the valence band is only half-filled as shown in Fig. 15.5a. As already discussed under metallic conductivity, the electrons here can respond to an externally applied field, by acquiring extra velocity and moving into the empty states in the top half of the valence band. In the second alternative, a full valence band overlaps the conduction band as shown in Fig. 15.5b so that there is no forbidden gap.

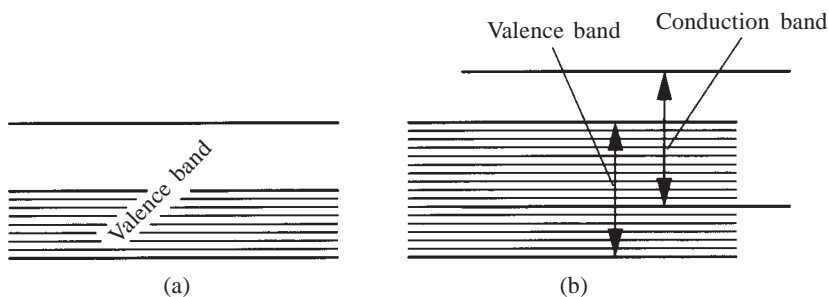


Fig. 15.5 Metals have partially filled or overlapping bands.

Monovalent metals such as Cu, Ag and Au have one electron in the outermost orbital. Correspondingly, the outermost energy band is only half full in these metals. Divalent metals such as Mg and Be have overlapping conduction and valence bands. Therefore, they also conduct even if the valence band is full. The band structure of trivalent metals such as Al is similar to that of monovalent metals, in that the outermost band is half full.

In elements of the fourth column, the electrons in the outermost orbital are even in number, as in the divalent metals. The valence band is full, but there is no overlap of the valence band with the conduction band here. Taking the case of covalent *diamond*, the potential energy of electrons is a strong function of their location in the crystal. Correspondingly, a forbidden gap exists and it is relatively large in magnitude (5.4 eV), see Fig. 15.6a. As we move down the fourth column to the elements below diamond, the electrons of the outermost orbital are farther away from the nucleus. The effect of this increasing separation is dominant over the effect of the increasing charge on the nucleus, so that the

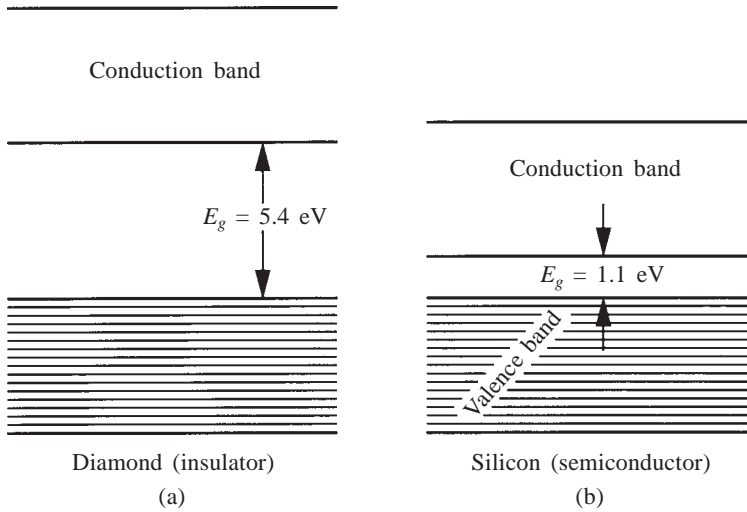


Fig. 15.6 The difference between (a) an insulator and (b) a semiconductor is in the magnitude of the energy gap.

electrons are less tightly bound to the nucleus, as we go down the column. The potential energy of the electron is no longer a strong function of its location. This reduces the energy gap from 5.4 eV in diamond to 1.1 eV in *silicon* (Fig. 15.6b), 0.7 eV in germanium, and a mere 0.08 eV in gray tin. In lead, the forbidden gap is zero.

Those materials which have an energy gap of about 2–3 eV or less are called *semiconductors* and those with a gap of more than 3 eV are known as *insulators*. In contrast to metals, both insulators and semiconductors have a finite forbidden energy gap, the only difference between them arising from the magnitude of the energy gap.

15.2 Intrinsic Semiconductors

In Chapter 14, we discussed the mechanism of conduction in metals. When an external electric field is applied, the free electrons accelerate by moving into the vacant quantum states immediately above the Fermi level. In semiconductors and insulators, this is not possible, as there is a forbidden gap.

In order to conduct, the electrons from the top of the full valence band have to move into the conduction band, *by crossing the forbidden gap*. The field that needs to be applied to do this works out to be extremely large. Take the example of silicon, where the forbidden gap is about 1 eV. This is approximately the energy difference between a location close to an ion core and another location away from the ion core. The distance between these two locations is about 1 \AA (10^{-10} m). Therefore, a field gradient of $\sim 1 \text{ V}/(10^{-10} \text{ m}) = 10^{10} \text{ V m}^{-1}$ is necessary to move an electron from the top of the valence band to the bottom of the conduction band. Such a high field gradient is not realizable in practice.

The other possibility by which this transition can be brought about is by *thermal excitation*. At room temperature, the thermal energy that is available can excite a limited number of electrons across the energy gap. This limited number accounts for semiconduction. When the energy gap is large as in diamond, the number of electrons that can be excited across the gap is extremely small. This accounts for the very high resistance of insulators, recall Table 14.1.

In intrinsic semiconductors, the conduction is due to the intrinsic processes characteristic of the crystal in question, *without the influence of impurities*. A pure crystal of silicon or germanium is an intrinsic semiconductor. The electrons that are excited from the top of the valence band to the bottom of the conduction band by thermal energy are responsible for conduction. The number of electrons excited across the gap can be calculated from the Fermi–Dirac probability distribution:

$$P(E) = \frac{1}{1 + \exp [(E - E_F)/kT]} \tag{15.2}$$

The Fermi level E_F for an intrinsic semiconductor lies midway in the forbidden gap, as illustrated in Fig. 15.7. The probability of finding an electron

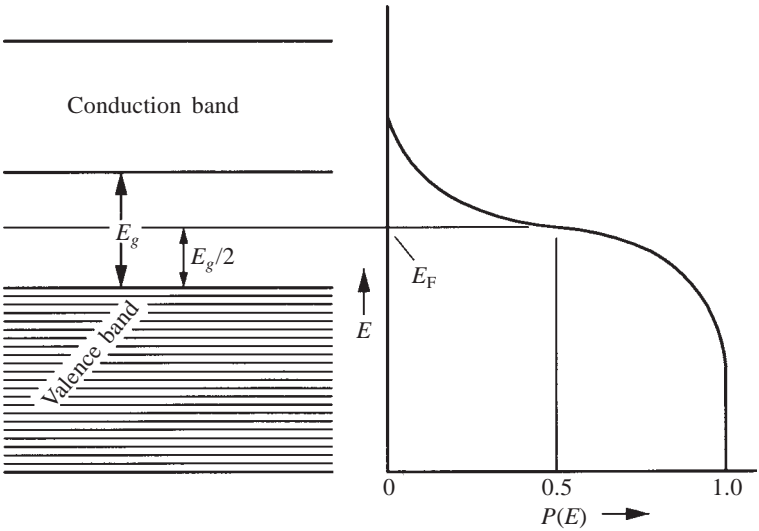


Fig. 15.7 The Fermi level in an intrinsic semiconductor lies in the middle of the energy gap.

here is 50%, even though energy levels at this point are forbidden. Then, $(E - E_F)$ in the denominator of Eq. (15.2) is equal to $E_g/2$, where E_g is the magnitude of the energy gap. For a typical semiconductor like silicon, $E_g = 1.1$ eV, so that $(E - E_F)$ is 0.55 eV, which is more than twenty times larger than the thermal energy kT at room temperature ($= 0.026$ eV). The factor unity in the denominator can therefore be ignored, in comparison to the exponential term, so that the probability $P(E)$ of an electron occupying energy level E becomes

$$P(E) = \exp \left[\frac{-E_g}{2kT} \right] \quad (15.3)$$

The fraction of electrons at energy E is equal to the probability $P(E)$. We can then write for the number n of electrons promoted across the gap:

$$n = N \exp \left[\frac{-E_g}{2kT} \right] \quad (15.4)$$

where N is the number of electrons available for excitation from the top of the valence band.

The promotion of some of the electrons across the gap leaves some vacant electron sites in the valence band. These are called *holes*. As each excited electron leaves back one hole, an intrinsic semiconductor contains an equal number of holes in the valence band and electrons in the conduction band, that is, $n_e = n_h$. Equation (15.4) gives the number of each of these species.

Under an externally applied field, the electrons, which are excited into the conduction band by thermal means, can accelerate using the vacant states available in the conduction band. At the same time, the holes in the valence band also move, but in a direction opposite to that of the electrons. The conductivity of the intrinsic semiconductor depends on the concentration of these charge carriers, n_e and n_h . In the case of metals, we referred to the drift velocity acquired by the free electrons in an applied field. Similarly, we can define the mobility of conduction electrons and holes μ_e and μ_h as the drift velocity acquired by them under unit field gradient.

$$\begin{aligned} \text{Mobility} &= \frac{\text{drift velocity}}{\text{field gradient}} \\ &= \frac{\text{m s}^{-1}}{\text{V m}^{-1}} \text{m}^2 \text{V}^{-1} \text{s}^{-1} \end{aligned}$$

We can then write the conductivity σ of a semiconductor as

$$\sigma = n_e e \mu_e + n_h e \mu_h \quad (15.5)$$

where e is the electronic charge and n_e and n_h are concentrations per unit volume. The mobilities of electrons and holes in silicon and germanium are listed in Table 15.1.

TABLE 15.1
Mobilities of Electrons and Holes in Silicon and Germanium

Species	Mobility at room temperature, $\text{m}^2 \text{V}^{-1} \text{s}^{-1}$	
	<i>Silicon</i>	<i>Germanium</i>
Electrons	0.14	0.39
Holes	0.05	0.19

Example 15.1 The resistivity of pure silicon at room temperature is 3000 ohm m. Calculate the intrinsic carrier density.

Solution The intrinsic charge carriers in pure silicon are electrons and holes in equal numbers. From Eq. (15.5), we obtain

$$\begin{aligned} n = n_e = n_h &= \frac{\sigma}{(\mu_e + \mu_h)e} \\ &= \frac{1}{(0.14 + 0.05) \times 3000 \times 1.602 \times 10^{-19}} \\ &= 1.095 \times 10^{16} \text{ m}^{-3} \end{aligned}$$

For an intrinsic semiconductor, the number of charge carriers (electrons and holes) is given by Eq. (15.4). This number is dependent on temperature in an exponential way and, therefore, increases very rapidly with increasing temperature. So, the conductivity also increases in a similar fashion with temperature. The small temperature dependence of mobilities can be neglected, as compared to the effect of temperature on the number of charge carriers. So, a plot of the logarithm of conductivity against the reciprocal of temperature (in kelvin) yields a *straight line*. The energy gap can be computed from the slope of this straight line.

15.3 Extrinsic Semiconductors

In extrinsic semiconductors, the conduction is due to the presence of *extraneous impurities*. The process of deliberate addition of controlled quantities of impurities to a pure semiconductor is called *doping*. The addition of impurities markedly increases the conductivity of a semiconductor. Consider a silicon crystal which is doped with a fifth column element such as P, As or Sb. As illustrated in Fig. 15.8a, the fifth column element, phosphorus in this case,

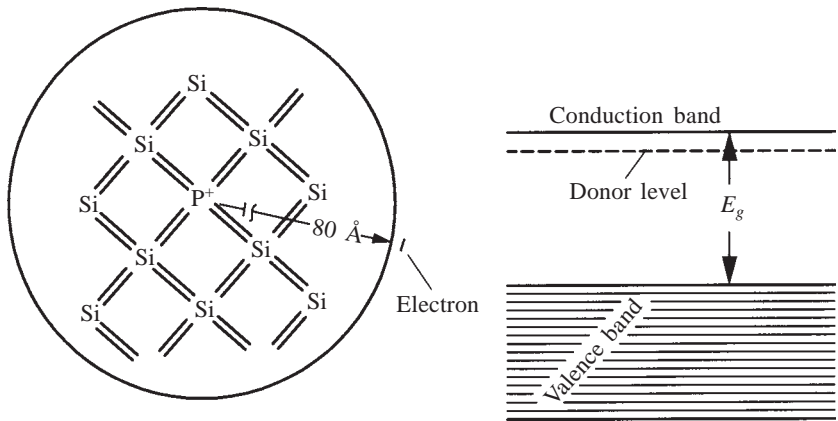


Fig. 15.8 Silicon doped with phosphorus from the fifth column of the periodic table becomes an *n*-type semiconductor.

substitutes for a silicon atom in the diamond cubic structure. Four of the five electrons in the outermost orbital of the phosphorus atom take part in the tetrahedral bonding with the four silicon neighbours. The *fifth electron* cannot take part in the discrete covalent bonding. It is *loosely bound* to the parent atom. It is possible to calculate an orbit for the fifth electron assuming that it revolves around the positively charged phosphorus ion, in the same way as for the 1s electron around the hydrogen nucleus. There is, however, one important difference. The electron of the phosphorus atom is moving *in the electric field of the silicon crystal* and not in free space, as is the case in the hydrogen atom. This brings in the dielectric constant of the crystal into the orbital calculations, and the radius of the electron orbit here turns out to be very large, about 80 Å as against 0.5 Å for the hydrogen orbit. Such a large orbit evidently means that the fifth electron is almost free and is at an energy level close to the conduction band, as shown in Fig. 15.8b.

Excitation of the fifth electron into the conduction band takes place much more readily than excitation from the valence band of the silicon crystal. The phosphorus atom is said to donate its electron to the semiconductor. The energy level of the fifth electron is called *the donor level*, see Fig. 15.8b. As the elements to the right of the fourth column donate negative charges (electrons), the semiconductors doped with them are called *n-type semiconductors*. The energy required to excite the fifth electron into the conduction band is known as the ionization energy, see Table 15.2.

TABLE 15.2

Ionization Energies for Some Elements in Silicon and Germanium (eV)

Impurity		Silicon	Germanium
<i>n</i> -type	Phosphorus	0.044	0.012
	Arsenic	0.049	0.013
	Antimony	0.039	0.010
<i>p</i> -type	Boron	0.045	0.010
	Aluminium	0.057	0.010
	Gallium	0.065	0.011
	Indium	0.16	0.011

As compared to the energy gap, the ionization energy of an impurity atom is very small. So, at room temperature, a large fraction of the donor level electrons are excited into the conduction band. This fraction is much larger than the fraction of electrons excited due to the intrinsic process, that is, from the valence band. According to the law of mass action, the product of the number of electrons in the conduction band and the number of holes in the valence band must be constant. This condition drastically reduces the number of holes in the *n*-type semiconductor. The electrons in the conduction band become the *majority charge carriers*.

Consider the alternative process of doping a silicon crystal with a third column element such as Al, Ga or In. Aluminium has three electrons in the outer orbital. While substituting for silicon in the crystal, it *needs an extra electron* to complete the tetrahedral arrangement of bonds around it. The extra electron can

come only from one of the neighbouring silicon atoms, thereby creating a vacant electron site (hole) on the silicon. The aluminium atom with the extra electron becomes a negative charge and the hole with a positive charge can be considered to revolve around the aluminium atom, leading to the same orbital calculations as above. This is illustrated in Fig. 15.9a.

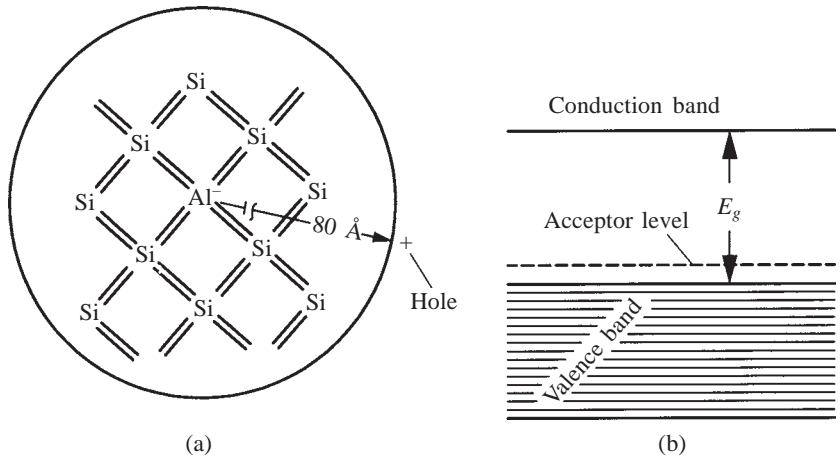


Fig. 15.9 Silicon doped with the third column element aluminium becomes a *p*-type semiconductor.

At 0 K, the holes remain bound to the impurity atoms. As the temperature is raised, the holes break away from the impurity atoms due to thermal excitation and become available for conduction. The ionization energy to free a hole bound to the impurity is approximately the same as the ionization energy of donor electrons in the same crystal, see Table 15.2. The bound-hole levels are called *acceptor levels* (aluminium accepts an electron) and are just above the valence band, as shown in Fig. 15.9b. The fraction of holes created by thermal excitation in doped silicon is much larger than those created by excitation of electrons into the conduction band. The law of mass action holds good here also and the positive holes become the majority charge carriers. Such an extrinsic semiconductor is called a *p*-type semiconductor.

Equation (15.5) describes the conductivity of extrinsic semiconductors as well. The only difference is that the number of electrons in the conduction band and the number of holes in the valence band are not equal in the case of an extrinsic semiconductor. One of the two dominates, depending on the type of the extrinsic process.

Example 15.2 As the concentration of electrons in a semiconductor is changed by changing the impurity level, the conductivity also changes. Show that it passes through a minimum when $n_e = n_i \sqrt{\mu_h / \mu_e}$ and find the minimum value. Here n_i is the intrinsic concentration.

Solution The conductivity of a semiconductor is given by

$$\sigma = n_e e \mu_e + n_h e \mu_h$$

Replacing n_h in terms of n_i and n_e ($n_i^2 = n_e n_h$), we have

$$\sigma = n_e e \mu_e + n_i^2 e \mu_h / n_e$$

Noting that n_i is a constant at constant temperature,

$$d\sigma/dn_e = e\mu_e - n_i^2 e \mu_h / n_e^2 = 0$$

for a minimum (the second derivative can be shown to be positive). So,

$$n_e = n_i \sqrt{\mu_h / \mu_e}$$

Also,

$$n_i^2 = n_e^2 \mu_e / \mu_h$$

Substituting this in the second equation above, we obtain

$$\sigma_{\min} = 2n_e e \mu_e$$

A typical plot of the logarithm of conductivity against temperature is shown for an extrinsic semiconductor in Fig. 15.10. The plot can be divided into three distinct regions. In the low temperature range (corresponding to large values of $1/T$), the conductivity is an exponential function of temperature. The electrons excited from the donor level to the conduction band (or the holes created by

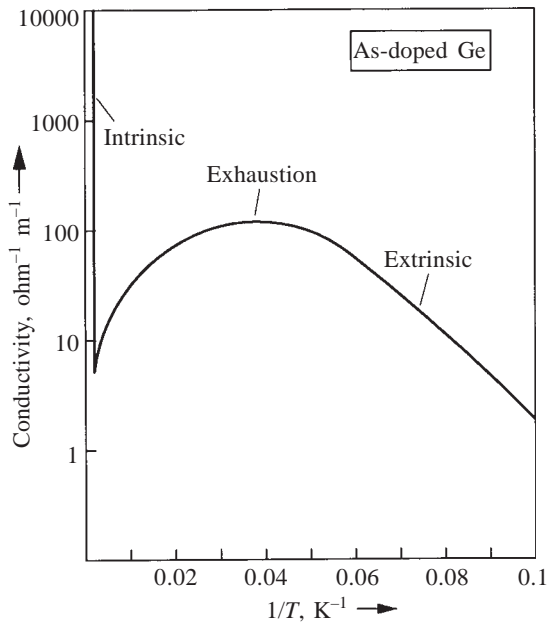


Fig. 15.10 The variation of log of conductivity with the reciprocal of temperature for an extrinsic semiconductor.

excitation to the acceptor level) become available in increasing numbers with increasing temperature. So, this region is called the *extrinsic region*. From the slope of the straight line in this region, the ionization energy of the impurity can be calculated. At higher temperatures, there is the second region called the *exhaustion region*. Here, the excitation of charge carriers due to impurities is nearing completion, due to the exhaustion of unexcited impurity electrons or holes. This is indicated by the flat region of almost zero slope. In this region, the slope can even become positive as in Fig. 15.10, due to the dominance of the temperature dependent mobility term in the conductivity equation. Increasing temperature results in decreasing mean free path of the conduction electrons or holes. So, the mobility decreases with increasing temperature. Above room temperature, there is another linear region with a negative slope of large magnitude, which corresponds to *intrinsic conduction*. The slope difference between the extrinsic and intrinsic regions reflects the difference in magnitude between the impurity ionization energy and the energy gap E_g .

For stability of operation of a semiconductor device, a relatively flat region, where the conductivity does not vary appreciably with changes in ambient temperature, is often desirable. On the other hand, if the device is to be used as a thermistor for measuring temperature, maximum sensitivity in measurement is desirable.

15.4 Semiconductor Materials

Silicon ($E_g = 1.1$ eV) is the most widely used semiconductor crystal. It is available in abundance in the earth's crust in the form of silica and silicates. It has a moderately high melting point (1410°C), which is easily achieved in modern zone refining and crystal growing apparatus. Germanium ($E_g = 0.7$ eV) is the other elemental semiconductor crystal with a lower melting point (937°C).

Apart from these, there are a number of compound semiconductors formed by combinations of equal atomic fractions of fifth- and third-column, or sixth and second-column elements. The crystal structures of these compounds are related to the diamond structure, as discussed in Chap. 5. Table 15.3 lists the properties of some III–V compound semiconductors.

TABLE 15.3
Properties of Some Semiconductor Compounds

Compound	Energy gap, eV	Mobility, $\text{m}^2 \text{V}^{-1} \text{s}^{-1}$		Melting point, °C
		<i>Electrons</i>	<i>Holes</i>	
GaP	2.26	–	0.002	1350
AlSb	1.52	0.02	0.02	1050
GaAs	1.43	0.85	0.04	1240
InP	1.29	0.46	0.015	1070
GaSb	0.78	0.4	0.07	705
InAs	0.35	2.3	0.024	940
InSb	0.18	6.5	0.1	525

We also have amorphous (noncrystalline) semiconductors. The fact that the local coordination and the interatomic distances characteristic of a crystal are largely unchanged in the noncrystalline state makes the band model to be approximately valid for the noncrystalline state as well.

15.5 Fabrication of Integrated Circuits

Si, Ge and GaAs are among the most important device materials. In all VLSI (very large scale integration) circuits, silicon is *the* material and has edged out Ge, which played an important role in the sixties. The advantages of Si over Ge in terms of cost and properties are compared below.

Property	Si	Ge
Energy gap	1.1 eV	0.66 eV
Upper temperature limit	150°C	100°C
Junction leakage current	less	more
Breakdown strength	higher	lower
Oxide quality	excellent	water soluble and unsuitable
Relative cost of electronic grade	1	10

GaAs and its ternary and quaternary derivatives are used in optoelectronic devices, as these compounds have a direct energy gap (the direction of motion of an electron during a transition across the energy gap remains unchanged). A direct gap is necessary for efficient optoelectronic conversion. A combination of Si and GaAs devices in integrated circuits is being developed. At present, over 98% of all devices are based on silicon only.

15.5.1 Production of Metallurgical Grade Silicon

The starting material is pure sand, which is available in plenty on earth's crust. Sand (SiO_2) is heated with carbon in an electric furnace to reduce it according to the reaction:

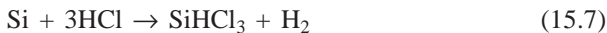


The silicon thus obtained is of 99% purity and is called the metallurgical grade silicon. This has to be purified further to a very low level of impurity content to make it suitable for use in devices.

15.5.2 Semiconductor Grade Silicon

The semiconductor grade silicon has only a few parts per billion of impurities. Starting from the metallurgical grade, it can be produced by the zone refining process as discussed in Sec. 7.7. The other common method is a chemical process.

The metallurgical grade silicon is dissolved in HCl:



The product trichlorosilane (SiHCl_3) is a liquid at room temperature. The fractional distillation of this liquid removes chlorides of dopants and of other impurities such as iron and copper and also SiCl_4 . A mixture of the purified SiHCl_3 and H_2 is then evaporated and passed through a reactor, which contains “slim rods” of high purity silicon. The gaseous mixture now undergoes the reverse reaction:



Solid silicon is deposited on the heated ‘slim rods’, which grow radially. This process can produce rods of semiconductor-grade, polycrystalline silicon up to 200 mm in diameter and 2–3 m long.

15.5.3 Single Crystal Growth

Single crystals of semiconductor grade silicon are grown either by the Czochralski (CZ) method or by the float zone (FZ) method. The Czochralski process (briefly referred to in Sec. 9.6) consists of a furnace with a gradient in temperature, as shown in Fig. 15.11. A typical CZ apparatus weighs about 20

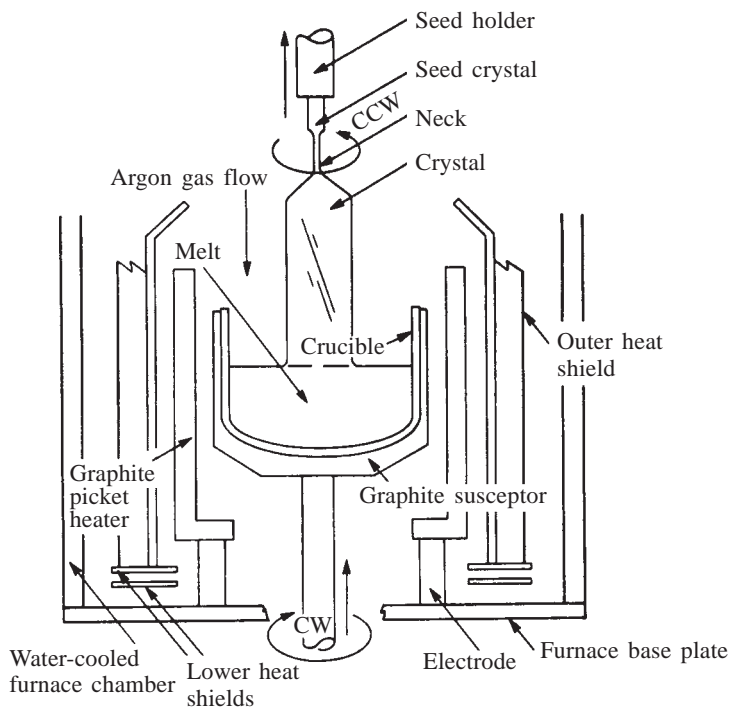


Fig. 15.11 The Czochralski (CZ) process of single crystal growth.

tonnes, is 7 to 8 metres tall and holds a charge of 100 kg of silicon. The main parts are the crucible, the susceptor, the heating element, power supply, seed shaft, rotation mechanisms, doping gas source, vacuum systems and process control through a microprocessor with sensors.

The crucible should have a melting point higher than that of silicon. High melting point carbide crucibles such as TiC and TaC are unsuitable, as they introduce metallic impurities. The best crucible material is found to be high purity SiO₂. The crucible slowly dissolves and introduces oxygen into the melt. The susceptor is made of nuclear-grade graphite and is the source of carbon contamination. Resistance heating is mostly used. The ambient is usually vacuum, which prevents graphite oxidation and removes oxygen continuously as silicon monoxide (SiO) vapour. Oxygen is interstitial in silicon and is electrically inactive. However, on supersaturation during cooling, it precipitates as SiO₂. The stresses generated thereby can induce defects like dislocations in the crystal. The carbon picked up from the susceptor is a substitutional impurity but is electrically inactive. It can be indirectly harmful, as it promotes the precipitation of oxygen.

The semiconductor grade polycrystalline silicon is melted in the silica crucible. A seed crystal of the desired orientation is made to touch the surface of the melt. Two commonly used orientations for the axis of the crystal are $\langle 111 \rangle$ and $\langle 100 \rangle$. The handle holding the seed is slowly pulled upwards at a speed of 50 to 100 mm per hour. The temperature T is maintained at $T = T_m$ at the liquid-seed interface, at $T < T_m$ above the interface and at $T > T_m$ below (inside the liquid). This avoids any nucleation of the solid crystal in the liquid and also prevents remelting of the growing crystal. The handle to which the seed is attached is rotated at 6 to 8 rpm and the crucible holding the liquid is rotated in the opposite direction at 10 to 12 rpm. These rotations ensure uniformity in temperature and avoid the appearance of hot or cold spots. By a suitable device, the crucible is gradually lifted upwards to compensate for the falling level of the liquid as the crystal grows. The crystal grows in the same orientation as the seed, without the need for the formation of a grain boundary at the interface. Crystal diameters up to 150–200 mm are possible.

If desired, the melt in the crucible of the CZ process can be doped with a suitable dopant. The concentration of the dopant in the crystal and the liquid that are in contact at the interface will in general be different from each other and will be according to the tie-line rule, see Sec. 7.3. The ratio of the concentration of the impurity (or dopant) in the solid c_s to that in the liquid c_l is known as the segregation coefficient k . The k values listed below for common impurities (except oxygen) are less than unity.

Al	As	B	C	Cu	Fe	P	Sb	O
0.002	0.3	0.80	0.07	0.0004	0.00001	0.35	0.023	1.25

This means that most impurities are left behind in the liquid. As the crystal grows, the melt becomes progressively enriched in the impurity (or dopant). A concentration gradient of the impurity thus develops along the axis of the grown crystal in the CZ method.

In the float zone (FZ) method, there is little chance of oxygen contamination, as no crucible is involved. The molten zone is held in position by the surface tension forces. Also, the dopant can be introduced from the vapour phase. If the partial pressure of the dopant gas is maintained constant, a uniform concentration of the dopant all along the length of the rod can be obtained. The float zone method, however, restricts the size of the crystal. Diameters larger than 75 mm are prone to zone collapse. To ensure that the crystal grows in the desired orientation, a seed is inserted at one end of the rod. The initial molten zone should be centred around the interface between the seed and the rod.

The grown crystals are generally free of edge dislocations, but may contain small dislocation loops arising from the condensation of supersaturated vacancies. Dislocations may also be introduced by thermal stresses during processing. Dislocations may play the following roles:

- (i) They act as nucleation sites for precipitation of impurities, e.g., SiO_2 from oxygen and AlFeSi from Al and Fe. The precipitates can in turn generate dislocations and stacking faults due to stresses from volume change and lattice mismatch at the interface.
- (ii) Dislocations can act as sinks for metallic impurities. This effect is beneficial in the 'gettering' process.
- (iii) They can alter diffusion profiles by providing high diffusivity paths.

Many thousands of tonnes of silicon single crystals are grown annually. They get converted into trillions of chips. A typical chip contains a million devices. A microprocessor can be built into a single chip. 85% of the single crystals are grown by the CZ method. Typical oxygen content in CZ crystals are 10^{18} atoms/cm³. CZ crystals are more suitable for highly-integrated low-power devices. 15% of the crystals are grown in the zone refining apparatus by the FZ method. Residual oxygen is about two orders of magnitude lower here. These crystals are suitable for less-integrated or discrete high-power devices.

15.5.4 Wafer Manufacture

The seed and defective-end portions are removed from the grown crystal. As precise control of diameter is not possible during growth, the crystal is ground to be perfectly round. Before cutting the crystal into thin discs called *wafers*, the orientation of the crystal should be determined. The wafer plane will have the same Miller indices as the axis of the crystal, as it is perpendicular to it. If it is (100) plane, it is necessary to locate the trace of the (010) or (001) plane on the wafer plane. This is done accurately by a single crystal x-ray method. A flat known as the *primary flat* ground on the surface of the crystal locates the crystal orientation. It is used later to align precisely the circuits to be printed on the wafer surface with respect to the cleavage planes (planes of easy fracture). A secondary flat may be ground to indicate whether the grown crystal is *n*-type or *p*-type.

The crystal is then sawed into thin slices of about 100 μm thick wafers. The inner diameter (ID) slicing is a common method. ID slicing uses a saw blade whose cutting edge is the interior of an annulus. The saw blade is a thin sheet of

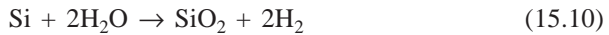
stainless steel with the inner rim impregnated with diamond powder. The wafer is lapped and the edges are rounded off to avoid breakage of the brittle silicon at points of stress concentration. To remove the damage and contamination during the shaping operations, etching is done in acid tanks with mixtures of HF, HNO₃ and acetic acid. Etching removes about 20 µm of the surface. Finally, polishing is done to a mirror-like finish on wheels suspended with fine alumina (Al₂O₃) powder.

15.5.5 Oxidation

Silicon has the unique ability to be oxidized into silica, which produces a chemically stable, protective and insulating layer on the surface of the wafer. The production of high quality ICs requires an understanding of the basic mechanisms of oxidation (see Sec. 13.1) and the ability to form in a controlled and repeatable fashion a high quality oxide. The functions of the oxide layer are to:

- (i) mask against diffusion or ion-implant;
- (ii) passivate the surface electrically and chemically;
- (iii) isolate one device from another; and
- (iv) act as a component in MOS devices.

Thermal oxidation is the principal technique and is carried out between 900 and 1300°C in dry oxygen or steam with the following reaction occurring:



During oxidation, the Si–SiO₂ interface moves into silicon. Experiments have established that oxidation proceeds by the diffusion of the oxidizing species (oxygen ion, oxygen atom or molecule) through the oxide layer to the Si–SiO₂ interface. The volume increases during oxidation. The constraint to match the oxide layer with the underlying silicon introduces stresses, which put the oxide layer in compression. If the stresses in the oxide layer were tensile in nature, the brittle layer will crack and will be unsuitable. The interface constraint also generates dislocations and stacking faults inside the silicon near the interface.

For long oxidation times, the oxidation rate is parabolic, i.e., the oxide thickness increases as the square root of time. As a function of temperature, the oxidation rate follows the Arrhenius law (recall Sec. 2.4) with a linear plot of \ln (oxidation rate) versus $1/T$. The activation energy for the process is found to be ~120 kJ mol⁻¹ for dry oxidation and ~70 kJ mol⁻¹ for wet (steam) oxidation.

If the silicon had been doped before oxidation, the dopant tends to redistribute itself in different concentrations in the two co-existing phases: Si and SiO₂. P, Ga, As and Sb dissolve more in Si than in silica; B behaves the opposite way. The segregation coefficient $k = c_{\text{Si}}/c_{\text{SiO}_2}$ is between 10 and 20 for P, Ga, As and Sb and less than one for B.

15.5.6 Photolithography

Lithography, as used in the manufacture of ICs, is the process of transferring geometrical shapes on a mask to the surface of a silicon wafer. The mask is to be prepared first. An electron beam machine generates the design pattern (reduced in size) and transfers it to a photosensitive glass called the *photomask* on a small area of a few mm². The photographic set-up is then moved to the adjacent area and the glass is exposed again, and so on, in a step-and-repeat fashion. As many identical areas are put on the mask as the number of chips that will ultimately fit on the wafer.

The photomask is positioned above the oxidized surface of the wafer, the wafer surface is coated with a photoresist solution and is exposed to an intense source of radiation, as shown in Fig. 15.12a. Till recently, UV light from a mercury lamp was used. Now, excimer lasers with wavelength in the range of 2480–1930 Å provide radiation of adequate intensity. (Excimer or excited dimer means a molecule consisting of an excited atom and an atom in its unexcited ground state.) After the exposure, when the photoresist is developed, it dissolves away from those areas where the light has fallen, as illustrated in Fig. 15.12b. In the next step of etching, see Fig. 15.12c, the etchant dissolves the oxide layer in the exposed areas only. The photoresist is etch-resistant. After etching, the left-over layer of photoresist is removed and the geometrical pattern is left behind in the silica, see Fig. 15.12d. All the steps in the image transfer are thus complete.

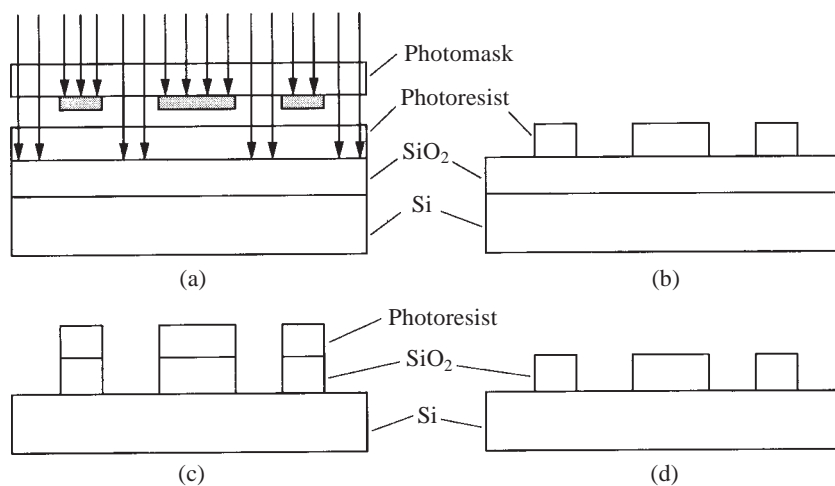


Fig. 15.12 The photolithographic process.

Photoresists can be positive or negative. Positive resists behave as described above, i.e., they dissolve away from areas where the light has fallen. Negative resists act in the opposite way. For example, polyisoprene combined with a photosensitive compound is a negative resist. Once activated by the light source, the compound transfers the absorbed energy to the polyisoprene chain molecules and cross-links them. The cross-linked polymer is insoluble in the developer.

The etching or window cutting can be through an acid medium (wet chemical etching) or by sputtering (dry etching). In reactive sputtering, a mixture of CF_4 and O_2 is used in the discharge tube to generate a plasma, which is weakly ionized gas mixture. When the plasma is directed onto the wafer surface, the ions collide with atoms in the exposed areas of the wafer surface, transfer their kinetic energy and eject the atoms out of the surface. The ions arrive mostly at normal incidence and the degree of etch anisotropy is high as compared to acid etching, i.e., the difference in the lateral dimension between the etch image and the mask image is small. The etching is also required to be selective, i.e., it should not remove the mask material or the underlying substrate, in which the previously processed device elements may be present.

15.5.7 Doping

Impurity doping can be carried out by solid state diffusion under two conditions: (i) a constant surface concentration of the dopant is maintained at a high enough temperature so that the dopant diffuses in and creates a characteristic concentration profile; and (ii) a constant total quantity of the dopant is first deposited on the surface (predeposition) either by short-time annealing or by other techniques such as ion implantation at a low temperature. The concentration profile is then altered through a high temperature anneal in a non-doping atmosphere (drive-in).

A p - n junction is created by diffusing into the bulk semiconductor through a window in the oxide layer an impurity, say, a p -type, where the bulk crystal has already been doped to be the n -type. In practice, the concentration-distance profile of the dopant is approximated to two limiting cases: the abrupt junction profile. This is the case in a shallowly-diffused junction or ion-implanted junction. In the linearly-graded junction, the concentration profile is assumed to change linearly with distance.

When the impurity is diffused in through a window, it moves downwards as well as sideways. The final shape of the diffused region, instead of being a flat box, becomes a box with bulging sides and spherical corners. This shape has an important effect on the junction breakdown characteristics.

Avalanche multiplication is the most important mechanism of junction breakdown. The avalanche breakdown voltage imposes an upper limit on the reverse bias of diodes, the collector voltage of bipolar transistors and on the drain voltage of MOSFETs. If the electric field is high, the carriers may carry enough energy such that their collisions with atoms generate electron-hole pairs. The newly-created pairs in turn generate more pairs by collisions and the process multiplies like an avalanche, eventually causing breakdown. The avalanche breakdown voltage decreases with increasing impurity concentration. It is also less for cylindrical shape of the sides of the diffused region and when the corners are spherical.

The diffusion parameters for common dopants are:

	B	P	As	Sb
D_0 (cm^2/sec)	0.76	3.85	24	0.214
Q (eV)	3.46	3.66	4.08	3.65

The diffusion coefficients of the above dopants in SiO_2 are several orders of magnitude smaller than for Si. Hence for all practical purposes, the oxide acts as a mask.

15.5.8 Ion Implantation

For doping at lower temperatures than those used in diffusion doping, ion implantation is used. Here, ionized-projectile atoms are introduced into the wafer with enough energy to penetrate the surface. Ion implantation is widely used for shallow doping of n^+ regions in the n -channel MOSFET and the base region of a bipolar transistor. For B, P and As dopant ions, an energy range of 3–500 keV is sufficient to implant them at a depth of 100–10,000 Å below the surface. The depth of penetration is approximately proportional to the ion energy. Precise control of depth and concentration of the dopant is possible. The ions produced by a source such as a hot cathode or arc discharge are accelerated under the potential and are projected onto the wafer, which is in electrical contact with a target holder which in turn is connected to a charge integrator. Electrons from the integrator neutralize the implanted ions as they come to rest in the wafer. Photoresists, oxides, nitrides, etc. act as masks during ion implantation. After implantation, a recovery anneal (see Sec. 9.8) is necessary to remove the excess point imperfections generated by the collisions of the ions with the atoms of the crystal. The recovery anneal is done by using a laser beam for local heating.

An interesting application of ion implantation is the introduction of high doses of ionized argon gas through the backside of the wafer. These generate a high density of crystal imperfections beneath the back surface. The damaged structure can capture unwanted diffusing impurities such as Cu, Fe, Au, etc. The process is called *gettering*. The damaged region with a high density of imperfections may undergo recrystallization when the wafer is heated, yielding a fine grain size locally. The grain boundaries thus generated also assist in gettering.

15.5.9 Epitaxial Growth

Epitaxial growth refers to growing of a crystal, using another crystal as the substrate, such that the atomic arrangement is continuous across the interface without the formation of a grain boundary. Homoepitaxy refers to silicon grown on silicon; the substrate and the grown crystal usually differ in the level and type of doping. Heteroepitaxy refers to a substantially different composition grown on the substrate. Examples are $\text{Al}_x\text{Ga}_{1-x}\text{As}$ grown on GaAs or a Si–Ge solid solution grown on Si. The two common processes of epitaxial growth are the chemical vapour deposition (CVD) and the molecular beam epitaxy (MBE).

In CVD of silicon, SiCl_4 is commonly used as the source. The following reaction occurs on the substrate surface:



The rate of deposition may be controlled by the rate of the above reaction and/or the rate of arrival/diffusing away of the reactants/products at the reaction

site. Hydrides of dopants (AsH_3 , PH_3 , etc.) are used as the source for simultaneous doping of the growing layer. Growth rates are in the range of 0.2–2 $\mu\text{m}/\text{min}$. A perfectly clean surface free of native oxide is required for epitaxial growth, as atom-to-atom contact is to be established. Etching with anhydrous HCl at 1200°C is done to clean the surface. The temperature for the reaction (15.11) is in the range of 1100 – 1200°C . This relatively high temperature may lead to autodoping, which is the diffusion of the dopant already present in the substrate into the growing layer through solid state diffusion or by diffusing out into the vapour phase and reentering. This can result in unintentional changes in dopant concentration profile.

In MBE, the temperatures used are lower: 400 to 800°C . Here autodoping is minimized. There are no chemical reactions in MBE. In high vacuum, evaporated silicon and the dopant are transported at high velocities to the substrate. The relatively low temperature ensures the condensation of silicon and the dopant on the substrate. Growth rates are in the range of 0.01 to $0.3 \mu\text{m}/\text{min}$.

Non-epitaxial films may be deposited during IC fabrication for the purpose of electrical insulation or chemical protection from the environment. Polycrystalline Si , SiO_2 and Si_3N_4 are usually deposited by the CVD process. Si_3N_4 acts as a very good barrier to the diffusion of water and sodium. Sodium from human sweat is known to enter the substrate and act as a fast-diffusing carrier resulting in high junction leakage currents.

15.5.10 Metallization

Metallization or routing is the process of providing electrical connections between different parts of the circuit. Aluminium is commonly used. It has a high electrical conductivity and a low melting point for easy evaporation during vacuum deposition. The vapour is produced by filament or electron-beam evaporation. It is transported in vacuum to the substrate. Co-evaporation of Al and Cu from an Al – Cu alloy results in a deposited Al layer with a small percentage of Cu in solution. The addition of Cu helps in reducing electromigration as discussed below.

Electromigration is the diffusion of an element under the influence of a current flow during the operation of a device. Electromigration in aluminium is a major problem in integrated circuits and is a common cause of failure of the circuit. Al diffuses along the grain boundaries and causes cavities to form at grain junctions, which coalesce and failure analogous to creep fracture ensues. The rate of electromigration of Al can be drastically reduced by adding Cu to the deposited Al . Cu acts by segregating to the grain boundaries of Al as a prestep to precipitation, recall Sec. 9.5. This segregation drastically reduces the diffusion coefficient of Al for grain boundary diffusion.

15.5.11 Circuit Simulation: Process simulation and integration

Several discrete steps in IC manufacture were outlined above. The sequence in which these steps are executed depends on the design features of the circuit to

be fabricated. In circuit simulation, the complete drawing of the circuit is broken into different levels of IC processing, e.g., gate electrodes on one level, contact windows on another level, and so on. These levels are called *masking levels*. The final IC is manufactured by sequentially transferring the features from each mask, level by level, to the wafer surface. Between two successive image transfers by photolithography, an ion-implant, doping, oxidation and/or metallization may take place.

After circuit simulation, it is also necessary to carry out a process simulation. Suppose a CMOS process consists of nine lithographic steps, six ion implantation and several diffusion, annealing and oxidation steps. The critical steps in the process are simulated on a computer, noting that all steps are strongly inter-related. Each of the thermal cycles in a process such as oxidation, epitaxial growth and annealing can affect the vertical and lateral diffusion of the dopants, interdiffusion between layers and so on. Developing a new process especially with a new material requires careful evaluation of the sequence of thermal cycles to be adopted.

Figure 15.13 illustrates a seven-mask process in building a circuit. The entire circuit is built in each of the squares of the wafer, which are separated

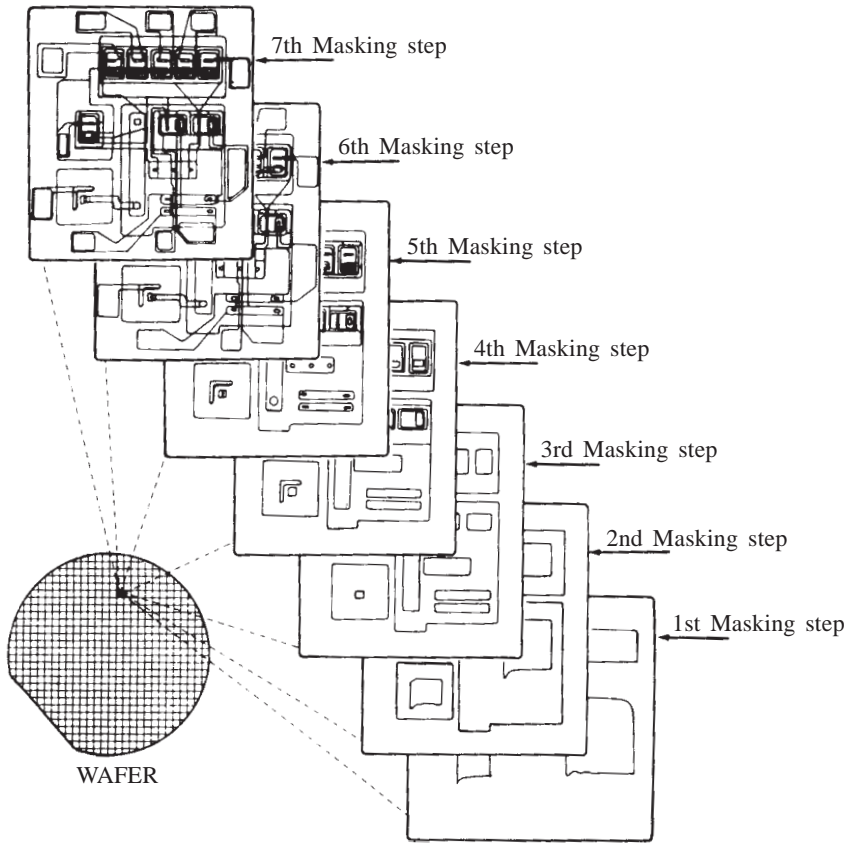


Fig. 15.13 The building of an integrated circuit with seven masking steps.

into chips later. The components of the circuit are arranged such that they occupy minimal space and interconnections and external connections are made as easy as possible. Miniaturization has been phenomenal in this process. We are now in the submicron technology, where the size of each component is 0.3–0.2 μm . The number of components per chip of a few mm^2 area was just 1 in 1960 and has now risen to several million. Currently, the most advanced circuits contain a billion devices, with the size of each being $\sim 1000 \text{ \AA}$ (0.1 μm).

After completion of all masking levels, the backside of the wafer may be just polished or a thin layer of gold may be vapour deposited on the backside, if the part is to be joined to another package later. The low melting temperature of the eutectic in the Au–Si phase diagram (363°C) enables such joining at a later stage at a relatively low temperature without causing any deterioration of the properties of the fabricated circuit.

The numerous chips on the wafer surface are separated by scribing and breaking either with a diamond-tip tool or a laser beam. The breaking is usually made to coincide with easy cleavage planes of Si and is done precisely along the boundaries of the repeating basic pattern. The chip then may be attached to a metallic or ceramic base. Al_2O_3 (alumina) is the most common ceramic base. Its high dielectric strength provides effective isolation between different parts of a complex package.

15.6 Some Semiconductor Devices

A number of semiconductor devices such as junction rectifiers, transistors, photocells, solar batteries and thermistors are known. They all use intrinsic or extrinsic semiconductor crystals. The same crystal consists of regions with different dopants giving them the *n*-type or the *p*-type characteristics. The boundary between two regions of opposite characteristics is called a *p-n* junction. It has special electrical properties such as the rectifying action. A three-region crystal (*p-n-p* or *n-p-n*) exhibits transistor action. Here, the superposition of the properties of the two junctions (*p-n* and *n-p*) close to each other acts as an amplifier of electric signals. The small size, reliability, low cost and low power consumption of the solid state devices have brought about a major revolution in the electronics industry in the last two decades.

15.6.1 Junction Diodes

The rectifying action of a *p-n* diode can be understood on the basis of the electronic structure of the semiconductor. When a pure semiconductor is doped to become *n*-type, the Fermi level shifts up from the middle of the energy gap towards the donor level. This is so because the position corresponding to 50% probability of occupation moves up due to the relatively high concentration of donor electrons in the conduction band. If the crystal is *p*-type, the Fermi level shifts down towards the acceptor level. When the same crystal is doped to become *n*-type on one side and *p*-type on the other side, the Fermi level has to

be constant throughout the crystal in thermal equilibrium. This results in the electron energy levels at the bottom of the conduction band in the n -part to be lower than those in the p -part, by an amount equal to the contact potential eV_0 , as shown in Fig. 15.14.

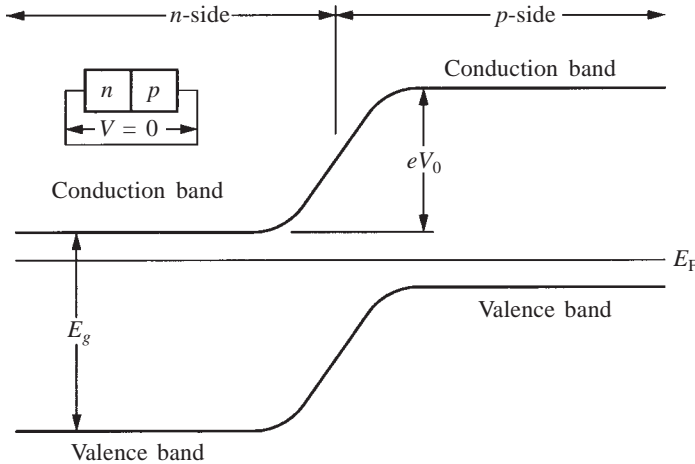


Fig. 15.14 The Fermi level is the same on both sides of a p - n junction. The contact potential at the junction gives rise to an energy barrier eV_0 .

At equilibrium, there is no net current flowing across the p - n junction. The concentration of electrons in the conduction band on the p -side is small. These electrons can accelerate down the potential hill across the junction to the n -side, resulting in a current I_0 which is proportional to their number. The concentration of electrons in the conduction band on the n -side is large in comparison, due to the donor contribution. However, only a small number of these electrons can flow to the p -side across the junction as they face a potential barrier. This small number can be computed using the Boltzmann probability equation. At equilibrium, the current from the n -side to the p -side and the current in the opposite direction, from the p -side to the n -side, are the same, equal to I_0 . There is an equal additional contribution to I_0 from the flow of holes across the junction. The concentration of holes is large in the p -region, as compared to the n -region.

If an external voltage V_i is now applied to the crystal such that the p -side becomes positive with respect to the n -side, the electron energy levels will change as shown in Fig. 15.15a. Note that the potential for electrons is opposite in sense to the conventional method of representing the sign of an electric potential. The barrier at the junction is now lowered by an amount eV_i , resulting in a greatly enhanced current flow in the forward direction, that is, from the n -side to the p -side. This change in barrier does not affect the flow of electrons in the reverse direction, from the p -side to the n -side, as the flow here is still *down* the potential hill. So, the applied voltage causes a large net current flow in the forward direction:

$$I_{\text{forward}} = I_0 \left[\exp \left(\frac{eV_i}{kT} \right) - 1 \right] \approx I_0 \exp \left(\frac{eV_i}{kT} \right) \quad (15.12)$$

The probability of jumps across a barrier is related to the height of the barrier through the Boltzmann probability factor. Here, the height of the barrier is lowered by an amount eV_i , which explains the positive sign inside the exponential in Eq. (15.12).

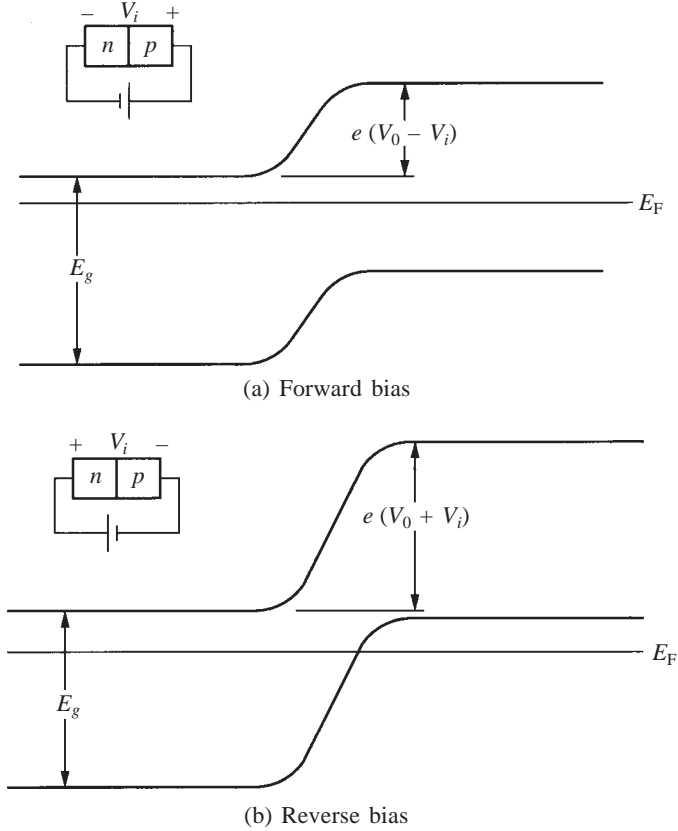


Fig. 15.15 The forward and the reverse bias of a p - n junction.

If an external voltage V_i is now applied in the reverse direction, the potential barrier for electrons at the junction is increased by an amount eV_i , as shown in Fig. 15.15b. This would drastically reduce the current flow from the n -side to the p -side. In the reverse direction, the flow remains unchanged at I_0 . The net current in the reverse direction from the p -side to the n -side is given by

$$I_{\text{reverse}} = I_0 \left[1 - \exp \left(-\frac{eV_i}{kT} \right) \right] \approx I_0 \quad (15.13)$$

As the barrier height is increased by an amount eV_i , there is the usual negative sign within the Boltzmann exponential in Eq. (15.13).

From Eq. (15.12), we note that the forward current increases exponentially. On the other hand, Eq. (15.13) indicates that the reverse current remains constant at a small value I_0 . This characteristic explains how a p - n junction can act as a rectifier. In Fig. 15.16, an alternating voltage is applied across the p - n junction. When the voltage is in the forward direction (called *forward bias*) in

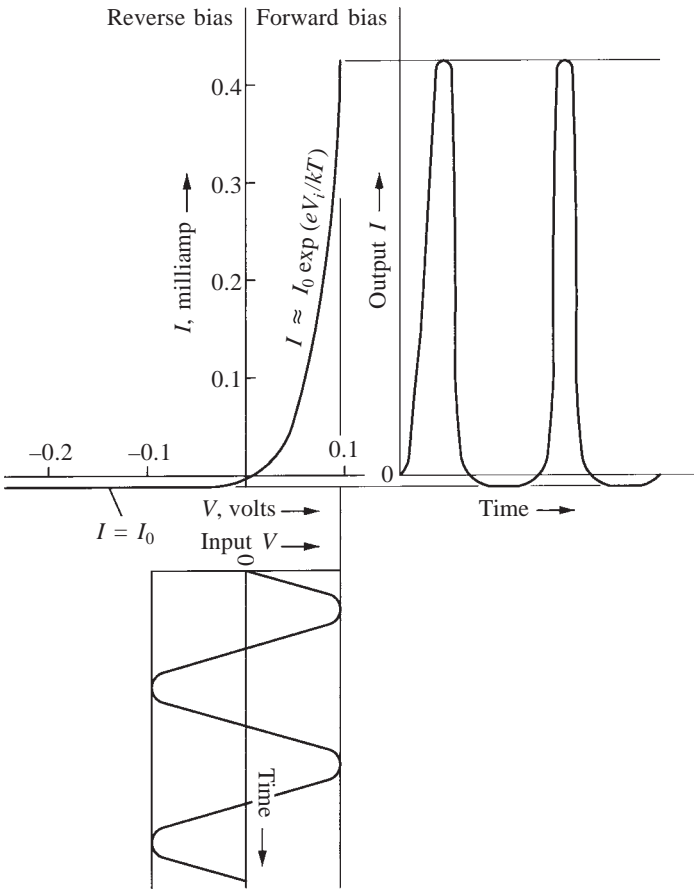


Fig. 15.16 Current-voltage characteristics of a p - n junction for forward and reverse bias.

one half of a cycle, the forward current increases exponentially with voltage. In the second half of the cycle, the voltage is in the reverse direction (*reverse bias*) and the current remains constant at the low equilibrium value of I_0 . The current output varies as a function time as indicated in the figure.

15.6.2 The Junction Transistor

The transistor is a device used to amplify an electric signal. It consists of two p - n junctions in the same crystal, arranged in the order n - p - n or p - n - p . In the n - p - n transistor, one of the two junctions is forward biased and connected to the

input signal. The n -region here is called the *emitter* and the p -region is the *base*. The other junction is reverse biased and connected to the output. The n -region at the output end is called the *collector*. Any input voltage V_i would reduce the contact potential at the emitter-base junction by an amount equal to eV_i . This would result in a large current flow from the emitter to the base region, as given by Eq. (15.12). The width of the base region is kept as small as possible so that the electrons flowing into the base region are not lost by recombination with the holes that are dominant in the base region. In an ideal transistor, almost all the electrons emitted reach the collector by flowing through the base and down the potential hill at the base-collector junction. A small change in the input voltage results in a large current from the emitter to the base and then on to the collector. The current flowing through different regions being the same, this large current can appear as a large voltage across the load resistor, if its resistance is high. Thus amplification results.

The above is a description of a simple n - p - n transistor. A number of more complex transistors are now known. Two promising types are MOSFET (metal-oxide-semiconductor field-effect transistor) and CCD (charge-coupled device). Details of the working of these devices are beyond the scope of this book. This is, however, not to minimize their practical importance. For example, a single computer now contains more than a million transistors (or one of its close relations MOSFET or CCD).

15.6.3 Junction Lasers

Laser stands for Light Amplification by Stimulated Emission of Radiation. When a p - n junction is forward biased, a large number of electrons flow across the junction from the n -side to the p -side. This greatly increases the concentration of electrons in the conduction band of the p -region, much above the equilibrium value. These electrons eventually recombine with the holes in the p -region, emitting in the process a coherent, monochromatic beam of light (the laser beam). The wavelength of this radiation for a GaAs junction laser is 8700 \AA . This light emerges strongly where the junction meets the sidewalls of the crystal. Mirrors are placed on opposite sides of the crystal to reflect the beam back and forth until it builds up in intensity. Alternatively, the flat boundary between the crystal and the air acts as a reflector.

As compared to silicon, the III-V semiconductor compound GaAs has the advantages of high signal speed, low power consumption and large operating temperature range. This combined with its laser properties makes it an ideal material for use in optical fibre and satellite communications and sophisticated supercomputers. The more recent development is the sandwich laser, in which a laser-active GaAs layer is sandwiched between two laser-inactive Ga-Al-As layers, one being p -type and the other n -type. The Ga-Al-As layers have larger energy gaps than GaAs and confine electrons to the active region. The excellent matching at the interface of the crystal structures of GaAs and Ga-Al-As (with less than 1% difference in the lattice parameter) ensures minimum strain and absence of plastic deformation during growth of the sandwich laser. The above

sophisticated applications require dislocation densities in the grown crystal of less than 10^7 m^{-2} . Also special growth techniques are needed, with precise control of the arsenic vapour pressure. As a semiconductor, GaAs is costly and will not replace silicon, which will continue to be used for making inexpensive chips.

If the fabrication difficulties associated with the 4% lattice mismatch between GaAs and Si is overcome, the monolithic GaAs/Si integration becomes possible. This will result in substantial improvements in very large scale integration (VLSI) performance, by adding the high-speed optoelectronic capability of GaAs to the conventional Si digital chips. For example, GaAs/Ga–Al–As optoelectronic interface units could provide high data-rate optical links to replace metallic interconnections between Si and VLSI subsystems.

A number of *p-n* junction lasers have been discovered in recent years extending the wavelength range from the far infrared to the ultraviolet region. The range of lasing wavelengths extends from 5900 Å for (GaIn)P to 300,000 Å for (PbSn)Te. Laser beams are used in a variety of applications such as carrier waves in communication engineering, fabrication of electronic devices, drilling holes in steel plates and for medical purposes, where a very localized heating is needed. As many as 40 different eye problems are treated now by lasers!

15.6.4 Photon Detectors

The electrons in the valence band can be excited into the conduction band by visible light, when the energy gap of a semiconductor is in the right range 1.60–3.20 eV. The additional holes and electrons created by the incident light can lead to an increase in the current in an appropriate external circuit. The current is a direct measure of the incident light intensity. *Photoconductors* are devices used for detecting and measuring light energy.

The sensitivity of a photoconductor is a maximum, if it remains essentially an insulator in the dark, that is, with a minimum of thermal excitation of charge carriers. Cadmium sulphide ($E_g = 2.42 \text{ eV}$) is a good photoconductor and responds to light in the green region. CdSe has a smaller gap of 1.74 eV, corresponding to the red end of the visible spectrum. CdTe with a gap of 1.45 eV is further out in the infrared region. PbS, PbSe and PbTe photoconductors have energy gaps in the range of 0.3 eV. Photoconductors are used as burglar alarms, for automatic door opening and for switching on street lights, as the sun goes down.

SUMMARY

1. Electrons travelling through the periodic array of atoms in a crystal are reflected when the Bragg equation is satisfied. The wave form of the electrons at this critical condition is two standing waves corresponding to two fixed electron locations of different potential energy. This gives rise to a break in the $E-k$ curve.

2. The energy gap in the $E-k$ curve occurs at different values of k for different crystal directions of electron motion. The net gap taking into account all directions of motion is called the forbidden gap. Electrons cannot take the energy values that lie in the forbidden gap.
3. Solids are classified as metals, semiconductors and insulators according to their band structure. Metals have partially filled bands or overlapping bands. Both insulators and semiconductors have a finite forbidden gap, the semiconductors having smaller gaps in the range of 1–3 eV.
4. The ability of a semiconductor or an insulator to conduct electricity depends on thermal excitation and the presence of impurities. In pure crystals, thermal excitation provides a small concentration of electrons in the unfilled conduction band, with an equal number of holes at the top of the valence band.
5. Impurities in an extrinsic semiconductor increase the concentration of charge carriers of either the negative type (donor electrons) or the positive type (acceptor holes).
6. The conductivity of a semiconductor is dependent on the number of charge carriers, their mobility and the electronic charge. The number of charge carriers is an exponential function of temperature, except in the exhaustion region of an extrinsic semiconductor.
7. Silicon is the most important elemental semiconductor. Compound semiconductors provide a range of energy gaps.
8. The fabrication of an integrated circuit comprises of a number of steps such as single crystal growth, wafer manufacture, oxidation, photolithography, doping, epitaxial growth and metallization.
9. A $p-n$ junction acts as a rectifier, allowing current to pass in one direction only. The $n-p-n$ junction is a transistor that amplifies electric signals. Complex transistors such as MOSFET and CCD perform a number of computer functions.
10. The III–V semiconductor compound GaAs has high signal speed and low power consumption as compared to silicon. Its laser function makes it ideally suited for use in satellite communications and supercomputers.

PROBLEMS

- 15.1 Explain why lead and zinc with an even number of electrons in the outer shell and a full valence band are conductors.
- 15.2 Estimate the temperature at which diamond will have the same conductivity as silicon at room temperature.

Answer: 1470 K.

- 15.3** Find the fraction of electrons excited into the conduction band in germanium at 100 K, 300 K and 1200 K.

Answer: 2.3×10^{-18} , 1.3×10^{-6} and 3.4×10^{-2} .

- 15.4** Why does the conductivity of a semiconductor change with impurity content? Compare this with the behaviour of metallic conductors.

- 15.5** The resistance of a sample of pure germanium has the following temperature dependence:

Temperature, K	312	354	385	420
Resistance, ohm	11.8	2.33	0.90	0.35

Determine the energy gap in germanium.

Answer: 0.78 eV.

- 15.6** Specify three elements that you would add to pure silicon to make it an extrinsic semiconductor of (i) the *n*-type, and (ii) the *p*-type.

- 15.7** The resistivity of intrinsic germanium at 27°C is 0.43 ohm m. Calculate the intrinsic carrier density at 27°C. Assuming that the number of electrons near the top of the valence band available for thermal excitation is $5 \times 10^{25} \text{ m}^{-3}$, calculate the energy gap for germanium.

Answer: $2.5 \times 10^{19} \text{ m}^{-3}$, 0.75 eV.

- 15.8** Indium phosphide (InP) has an energy gap of 1.29 eV. The electron and hole mobilities for this semiconducting compound are 0.46 and $0.015 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$, respectively at 300 K. Calculate its conductivity.

Answer: $\sim 5 \times 10^{-5} \text{ ohm}^{-1} \text{ m}^{-1}$.

- 15.9** The donor ionization energy in germanium doped with phosphorus is 0.012 eV. Calculate the fraction of donor electrons promoted to the conduction band at 27°C, assuming a Boltzmann distribution.

Answer: 0.63.

- 15.10** Calculate the fraction of holes present at 300 K in silicon doped with indium. The acceptor level is 0.16 eV above the top of the valence band.

Answer: 2×10^{-3} .

- 15.11** Germanium is doped with 10^{-3} atomic fraction of antimony. Calculate the electron and the hole densities in this extrinsic semiconductor. Assume that the electron density in the conduction band is entirely that due to donor excitation. Calculate the resistivity of this material at room temperature.

Answer: $3.0 \times 10^{25} \text{ m}^{-3}$; $2.1 \times 10^{13} \text{ m}^{-3}$; $5.3 \times 10^{-7} \text{ ohm m}$.

- 15.12** A rod of *p*-type germanium 10 mm long, and 1 mm dia has a resistance of 100 ohm. What is the concentration of the impurity in this rod?

Answer: $6.2 \times 10^{21} \text{ m}^{-3}$.

- 15.13** Calculate the drift velocity of an electron moving in a germanium crystal in a field of gradient 1 kV m^{-1} .
Answer: 390 m s^{-1} .
- 15.14** 5 g of *n*-type silicon doped with arsenic of conductivity $10 \text{ ohm}^{-1} \text{ m}^{-1}$ is melted with 5 g of *p*-type silicon doped with aluminium having the same conductivity. What is the conductivity of the resulting material?
Answer: $3.2 \text{ ohm}^{-1} \text{ m}^{-1}$.
- 15.15** The equilibrium current across an unbiased *p-n* junction is $10 \text{ }\mu\text{A}$. Calculate the current when the junction is (i) forward biased by 0.1 V , and (ii) reverse biased by 0.1 V .
Answer: $479 \text{ }\mu\text{A}$, $10 \text{ }\mu\text{A}$.
- 15.16** A transistor has a collector current of 5 mA , when the emitter voltage is 20 mV . At 30 mV , the current is 30 mA . Calculate the current, when the emitter voltage is 40 mV .
Answer: 180 mA .
- 15.17** A CdTe photoconductor responds to a glowing cigarette. Explain why this is so.
- 15.18** Determine the wavelength of the light from a GaAs laser, assuming that the radiative transition occurs between the bottom of the conduction band and the top of the valence band.
Answer: $8670 \text{ }\text{\AA}$.

MULTIPLE CHOICE QUESTIONS

- The first reflection of free electrons in a BCC crystal occurs at the following value of k
 A. $2/a$ B. $\pi a/\sqrt{2}$ C. π/a D. $\sqrt{2}\pi/a$
- If an electron is moving at 60° to the (100) plane in a cubic crystal of lattice parameter a , the wave number k at the critical diffracting condition is:
 A. π/a B. $2\pi/a$ C. $\sqrt{2}\pi/a$ D. $2/\sqrt{3} \times \pi/a$
- The first Brillouin zone in Cu is bounded by planes of the type
 A. 100, 110 B. 111 C. 111, 220 D. 111, 200
- The energy gap in diamond is
 A. 5.4 eV B. $2\text{--}3 \text{ eV}$ C. 1.1 eV D. 0.08 eV
- The field gradient (V m^{-1}) required to accelerate an electron in cubic diamond ($E_g = 5.4 \text{ eV}$; $a = 3.57 \text{ }\text{\AA}$) over a distance equal to the atomic radius is
 A. 7×10^{10} B. 3.5×10^{10} C. 5.4 D. 1.5×10^{10}

19. Tick the value of energy gap of Si in the following:
 A. 0.08 eV B. 0.66 eV C. 1.1 eV D. 5.4 eV
20. For silicon doped with B,
 A. $n_e \gg n_h$ B. $n_h \gg n_e$ C. $n_h \gg n_i$ D. $n_h = n_e$
21. Electron and hole mobilities in Si are 0.14 and $0.05 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. The electrical resistivity in ohm-m at 300 K of Si doped with B of concentration equal to 10^{26} m^{-3} is (ionization energy of B = 0.045 eV)
 A. 1.4×10^5 B. 2.55×10^{-6} C. 7.14×10^{-6} D. 1.25×10^{-6}
22. Electron and hole mobilities in Si are 0.14 and $0.05 \text{ m}^2 \text{ V}^{-1} \text{ s}^{-1}$. 10 g of *n*-type Si doped with As is melted with *x* g of *p*-type Si doped with Al, both of the same dopant concentration, to yield Si of intrinsic conductivity. *x* is equal to
 A. 28 B. 10 C. 3.6 D. 1.8
23. The temperature at which 20% of the donor electrons are excited into the conduction band in phosphorus-doped silicon (ionization energy = 0.044 eV) is
 A. 44°C B. -2°C C. 737°C D. 1022°C
24. Metallurgical grade Si has a purity of about
 A. 99% B. 99.99% C. 99.999% D. 99.9999%
25. During purification of Si, the liquid that is produced by dissolving silicon in HCl is
 A. SiCl_4 B. SiH_2Cl_2 C. SiHCl_3 D. SiH_4
26. Trichlorosilane is
 A. SiHCl_3 B. SiH_2Cl_2 C. SiH_3Cl D. SiCl_3
27. The quality of oxide on Ge is
 A. excellent B. satisfactory C. unsuitable D. water soluble
28. As compared to the CZ method of crystal growth, the FZ method has the following advantages:
 A. larger diameter crystals can be grown
 B. uniform doping is possible
 C. oxygen contamination does not occur
 D. incorporation of substitutional carbon is an advantage
29. To reduce the dopant concentration gradient along the grown crystal to a minimum in the CZ process, the ratio c_l/c_s (*c* is concentration of dopant in liquid or solid) should be
 A. very low B. very high C. 0 D. near unity
30. As compared to Si, the electron mobility in GaAs is
 A. slower by about five times B. same
 C. faster by about six times D. faster by about 200 times

31. The two common orientations during Si single crystal growth is
A. $\langle 111 \rangle \langle 100 \rangle$ B. $\langle 110 \rangle \langle 001 \rangle$
C. $\langle 111 \rangle \langle 101 \rangle$ D. $\langle 100 \rangle \langle 001 \rangle$
32. For lasing action, the energy gap of a semiconductor should be
A. direct gap B. indirect gap C. negative gap D. $\text{gap} \geq 5 \text{ eV}$
33. GaAs has an energy gap of 1.43 eV. The wavelength of the radiation emitted during an electronic transition in GaAs will be in the
A. visible range B. ultraviolet range
C. infrared region D. far ultraviolet region
34. In the Czochralski method, the temperature T at the liquid-seed crystal interface is
A. $= T_m$ B. $< T_m$ C. $> T_m$ D. none of these
35. If the activation energy for oxidation of Si is 120 kJ mol^{-1} , the ratio of oxidation rates at 1300 and 1100°C is
A. 3.8 B. 1.0 C. 7.5 D. 1.7
36. The thickness of the oxide layer grown on a wafer is 400 \AA after 4 min. After 16 min at the same temperature, it is
A. 6400 \AA B. 1600 \AA C. 800 \AA D. 400 \AA
37. The thickness of the SiO_2 layer grown on a Si wafer is 400 \AA after 5 min at 1100°C . The thickness will be double after
A. 10 min B. 20 min C. 7.07 min D. 3.54 min
38. The typical temperature used for oxidation of a Si wafer is
A. 400°C B. 700°C C. 1100°C D. 1400°C
39. The advantages of ion-implantation over diffusion doping are
A. it is a low temperature process
B. point imperfections are not produced
C. shallow doping is possible
D. gettering is possible
40. The following can be grown epitaxially on Si without creating significant distortion or imperfections
A. Si of a different doping B. SiO_2
C. GaAs D. none of these
41. The grown single crystal generally contains
A. tilt boundaries
B. dislocation loops due to vacancy condensation
C. twin boundaries
D. grain boundaries
42. The segregation coefficient, $k = c_s/c_l$, for oxygen in silicon
A. 0.00001 B. 0.0004 C. 0.0002 D. 1.25

43. The functions of an oxide layer during IC fabrication can be to
 A. mask against diffusion or ion-implant
 B. insulate the surface electrically
 C. produce a chemically stable surface
 D. increase the melting point of silicon
44. A negative resist
 A. becomes more soluble after UV exposure
 B. becomes less soluble after UV exposure
 C. gets cross-linked
 D. dissolves during acid etching
45. The main difficulty in monolithic integration of Si and GaAs is the difference in the lattice parameters, which is about
 A. 0.1% B. 4% C. 25% D. 100%
46. Electromigration in metallization refers to the diffusion (under the influence of current) of
 A. Al B. Cu in Al–Cu alloy
 C. Si D. Na
47. Electromigration in an IC chip refers to
 A. grain boundary diffusion of Al in Al interconnections
 B. grain boundary diffusion of Cu in Al interconnections
 C. diffusion of Si in Al interconnections
 D. diffusion of oxygen in Si
48. A transistor has a collector current of 5 mA, when the emitter voltage is 20 mV. At 30 mV, the current is 30 mA. At 50 mV, it is
 A. 80 mA B. 180 mA C. 480 mA D. 1080 mA

Answers

- | | | | | |
|-------|---------------|-------------|-------------|----------|
| 1. D | 2. D | 3. D | 4. A | 5. A |
| 6. D | 7. A, B, C, D | 8. C | 9. B | 10. D |
| 11. D | 12. C | 13. A | 14. A | 15. A, B |
| 16. C | 17. D | 18. A | 19. C | 20. B, C |
| 21. C | 22. B | 23. A | 24. A | 25. C |
| 26. A | 27. C, D | 28. B, C | 29. D | 30. C |
| 31. A | 32. A | 33. C | 34. A | 35. A |
| 36. C | 37. B | 38. C | 39. A, C, D | 40. A |
| 41. B | 42. D | 43. A, B, C | 44. B, C | 45. B |
| 46. A | 47. A | 48. D | | |

Sources for Experimental Data

- P. Gise and R. Blanchard, *Modern Semiconductor Fabrication Technology*, Prentice-Hall, Englewood Cliffs, NJ (1986).
- S.P. Keller (Ed.), *Handbook on Semiconductors, Vol. 3—Materials, Preparation and Properties*, North Holland, Amsterdam (1980).

Suggestions for Further Reading

- C. Kittel, *Introduction to Solid State Physics*, John Wiley, New York (1976). Chaps. 7 and 8.
- K.K. Ng, *Complete Guide to Semiconductor Devices*, IEEE Press and Wiley-Interscience, New York (2002).
- S.M. Sze, *Physics of Semiconductor Devices*, John Wiley, New York (1981).

16 Magnetic Materials

New and important applications based on the magnetic properties of materials have come into prominence in the last two decades, supplementing such traditional applications as in the making of an ordinary transformer core. Our understanding of the microstructural factors that influence the magnetic properties is now better than before. The control of microstructure for obtaining the desired magnetic properties is today almost as important as the control necessary for achieving the optimum mechanical properties.

We will first consider the terms and definitions used in magnetism, to be followed by a discussion of ferromagnetism and related phenomena. Ferromagnetic and ferrimagnetic materials are by far the most important from the point of view of practical applications. Soft and hard magnetic materials are discussed in the last two sections.

Units

Quantity	SI units		Other units
	Unit	Symbol	
Magnetic induction B	weber per square metre or tesla	Wb m^{-2}	gauss
Saturation induction B_s		or	
Residual induction B_r		T	
Magnetic field strength H	ampere per metre	A m^{-1}	oersted
Coercive field H_c			
Magnetization M			
Magnetic permeability μ	henry per metre	H m^{-1}	—
Relative permeability μ_r	—	—	—
Magnetic susceptibility χ	—	—	—
Magnetic moment m	ampere metre ²	A m^2	—
Energy product $B_r H_c$	joule per cubic metre	J m^{-3}	—
Hysteresis loss	watt	W	—
Eddy current loss			
Frequency of alternating current	hertz	Hz	per sec

Constants

Permeability of free space $\mu_0 = 4\pi \times 10^{-7} \text{ H m}^{-1}$	
	$= 1.257 \times 10^{-6} \text{ H m}^{-1}$
Bohr magneton μ_B	$= 9.273 \times 10^{-24} \text{ A m}^2$
(magnetic moment unit)	

16.1 Terminology and Classification

The magnetic induction or magnetic flux density B (expressed in units of tesla or weber per square metre) due to a magnetic field applied in vacuum is related to the field strength H (in units of A m^{-1}) as follows:

$$B = \mu_0 H \quad (16.1)$$

where μ_0 is called the permeability of free space (vacuum). The units of permeability are

$$\begin{aligned} \frac{B}{H} &= \frac{\text{Wb m}^{-2}}{\text{A m}^{-1}} \quad (\text{Wb m}^{-2} = \text{N A}^{-1} \text{ m}^{-1}) \\ &= \text{H m}^{-1} \quad (\text{H} = \text{Wb A}^{-1}) \end{aligned}$$

In SI units, the permeability of free space has a value of $4\pi \times 10^{-7} \text{ H m}^{-1}$. It was originally defined to have a value of unity in cgs units.

If the field is applied to a solid medium, the magnetic induction in the solid is given by a similar relationship

$$B = \mu H \quad (16.2)$$

where μ is the permeability of the solid material through which the magnetic lines of force pass. In general, μ is not equal to μ_0 . The ratio of μ/μ_0 is called the *relative permeability* μ_r of the solid. The magnetic induction in a solid is also defined by

$$B = \mu_0 (H + M) \quad (16.3)$$

where M is called the *magnetization* of the solid. The magnetization is directly related to the applied field H through the *susceptibility* of the medium χ :

$$\chi = \frac{M}{H} \quad (16.4)$$

The orbital motion of charge carrying electrons in an atom is analogous to a current carrying coil. When a magnetic field is applied to an atom, the motion of the orbital electrons gets modified in such a way that a weak magnetic moment opposing the field is induced. *Diamagnetism* is the result of this interaction. In a diamagnetic solid, the magnetic lines of force due to an applied field are repelled, as shown in Fig. 16.1a. The diamagnetic susceptibility is negative and very small in magnitude, typical values of the order of -10^{-5} . An important

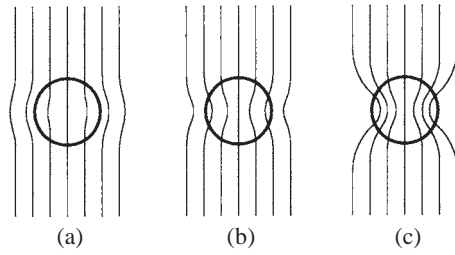


Fig. 16.1 (a) A diamagnetic solid repels slightly the magnetic lines of force. (b) A paramagnetic solid weakly attracts the lines of force. (c) A ferromagnetic solid attracts them very strongly.

exception, however, is a superconductor, which is perfectly diamagnetic with a susceptibility value of -1 . All the lines of force are repelled by the superconductor, thus making them useful for the purpose of shielding out magnetic fields.

Some atoms and molecules possess intrinsic permanent magnetic moments. In the absence of an externally applied field, the moments of the atoms in a solid are randomly oriented with respect to one another and the solid as a whole has no net magnetic moment. If an external field is applied, the magnetic moments tend to *align themselves parallel to the applied field*, so as to lower their potential energy. It is well known that a suspended magnetic needle aligns itself spontaneously with the earth's magnetic field. Similarly, there is a spontaneous tendency for the permanent moments of the atoms of the solid to align themselves in the direction of the field, thereby intensifying the lines of force in the field direction. This phenomenon is called *paramagnetism* and is illustrated in Fig. 16.1b. The aligning force on the permanent moments of the atoms with ordinary magnetic fields is rather small, so that the paramagnetic effect is weak. The paramagnetic susceptibility is *small and positive*, of the order of 10^{-3} . It decreases with increasing temperature, as thermal energy tends to randomize the alignment.

Ferromagnetic solids are those in which the permanent magnetic moments (due to electron spins) are already aligned due to bonding forces. The susceptibility is very large and positive for ferromagnetic materials, in the range 10^2 – 10^5 . They strongly attract the magnetic lines of force, as schematically shown in Fig. 16.1c.

16.2 Magnetic Moments due to Electron Spin

Permanent magnetic moments can arise from three sources: the orbital magnetic moment of the electrons corresponding to the quantum number m_l , the spin magnetic moment of electrons corresponding to the spin quantum number m_s , and the spin magnetic moment of the nucleus. Of these, the spin magnetic moments of the electrons are the only ones that are important from our point of view. We will discuss only this source of magnetism.

An electron can be crudely approximated to a charge-carrying particle spinning about its axis and creating its own magnetic field. The magnetic moment of an electron spin is taken as one unit, called the *Bohr magneton* μ_B . It has a value of $9.273 \times 10^{-24} \text{ A m}^2$. The net magnetic moment of two electrons of opposite spins is zero. Atoms or molecules that have quantum states all of which have paired electrons have zero net magnetic moment. A number of atoms and molecules, however, have unpaired electrons. For instance, the alkali metal atoms have only one electron in the outermost *s* orbital. This electron is unpaired and can align itself in an applied field giving rise to paramagnetism.

The order of filling of electron orbitals in atoms is given by *Hund's rule*. An atom with three electrons in a *p* orbital will have all three spins aligned, giving rise to a net magnetic moment of three Bohr magnetons. On the other hand, an atom with four electrons in the *p* orbital will have a net moment of only two units, as the spin of the fourth electron is opposite to that of the first three. Similarly, an atom with five electrons in a *d* orbital has a net magnetic moment of five units. But an atom with nine electrons in the *d* orbital has a net moment of only one. Seven electrons in an *f* orbital will give a net magnetic moment of seven units.

This simple deduction of the net magnetic moment of an atom has to be modified in the solid state, where the electron energy levels of the outermost orbitals interact and form energy bands. In the transition metals, the energy bands due to the *4s* orbital and the *3d* orbital may *overlap* to some extent. The *4s* electrons spend some time in the *3d* band. If the *3d* electrons are five or more in number, the spins of the overlapping *4s* electrons must be opposite to that of the unpaired *3d* electrons. Such an overlap, therefore, *reduces* the net magnetic moment. The iron atom with an outer electronic configuration of $3d^6 4s^2$ has a moment of 4 units as a free atom, but has a moment of only 2.2 as part of the crystal. Similarly, cobalt with a configuration of $3d^7 4s^2$ has a moment of 1.7 and nickel with a configuration of $3d^8 4s^2$ has a moment of only 0.6 unit in the crystal. In the case of gadolinium, the net magnetic moment of seven in the free atom remains unchanged in the solid state, as there is no overlap of the *4f* orbitals with the other energy bands.

The magnetization of a solid is the sum of the magnetic moments in unit volume of the solid.

$$\begin{aligned}\text{Magnetization} &= \text{magnetic moments/volume} \\ &= \text{A m}^2/\text{m}^3 \\ &= \text{A m}^{-1}\end{aligned}$$

Example 16.1 The saturation magnetization of BCC iron is 1750 kA m^{-1} . Calculate the net magnetic moment per iron atom in the crystal.

Solution

$$\begin{aligned}
 \text{The lattice parameter of BCC iron} &= 2.87 \text{ \AA} \\
 \text{Volume of the unit cell} &= 2.87^3 \times 10^{-30} \text{ m}^3 \\
 \text{Number of atoms in the unit cell} &= 2 \\
 \text{Net magnetic moment per atom} &= 1750 \times 1000 \times 2.87^3 \times 10^{-30} \times 1/2 \\
 &= 2.068 \times 10^{-23} \text{ A m}^2 \\
 \text{In units of } \mu_B, \text{ the moment} &= 2.068 \times 10^{-23} / (9.273 \times 10^{-24}) \\
 &= 2.2
 \end{aligned}$$

16.3 Ferromagnetism and Related Phenomena

Hund's rule is a manifestation of the spin dependent electrostatic energy of the electrons in an orbital. When the electrons have the same momentum and the same spin, a certain distance must separate them from one another, in order to be consistent with the Pauli exclusion principle. This physical separation tends to increase their kinetic energy, but reduces the electrostatic repulsive energy between them. Whether the electrons will align their spins or not will depend on the sign of the net change in energy. The stable, energetically favoured order of filling of electronic orbitals in atoms is given by Hund's rule.

In the solid state, the outer electronic orbitals of neighbouring atoms overlap and produce energy bands. Consider the overlapping and the formation of energy bands in the first transition metals. At the left end of the series, the radius of the 3d orbitals in the atoms is large enough to result in good overlap between neighbours in the crystalline state. The 3d band then contains all paired up electrons and there is no net magnetic moment in the crystal. We can say that the atoms are *antiferromagnetically coupled*, as the magnetic moments of a pair of atoms exactly cancel out.

As we go to the right in the transition series, the 3d orbitals *shrink* due to greater attraction from the increasing charge on the nucleus and consequently the overlap of the 3d orbitals decreases. The elements Fe, Co and Ni are in special situation. The unpaired electrons in the 3d orbitals of neighbouring atoms *align their spins* in a parallel fashion and thereby lower their spin dependent electrostatic energy. This lowering is partly offset by the rise in the Fermi level and the consequent increase in the average kinetic energy of the electrons. The increase in the Fermi level is directly attributable to the greater physical separation between electrons. The net gain in energy, $E_{\text{unmagnetized}} - E_{\text{magnetized}}$, called the exchange interaction energy is a *sensitive function of the ratio of the atomic diameter to the 3d orbital diameter*. This ratio for some of the first transition elements is shown as follows:

Element	Ti	Cr	Mn	Fe	Co	Ni
$d_{\text{atomic}}/d_{3d \text{ orbital}}$	1.12	1.18	1.47	1.63	1.82	1.98

It turns out that, *only when this ratio lies between 1.5 and 2.0, the exchange energy is positive* and parallel spins are energetically favoured. Among the common metals, *only Fe, Co and Ni have positive exchange energy and are in the spontaneously magnetized state*. These metals or their ions form the basis of the majority of magnetic materials.

No overlap of the d orbitals occurs in the higher transition series, as the d electrons are strongly attracted to their respective nuclei, which have larger positive charges. The only other elemental crystals where the exchange energy is appreciable belong to the first rare earth series such as gadolinium, terbium and dysprosium. Here, the $4f$ electrons align themselves in a parallel fashion.

Thermal energy tends to randomize the aligned spins, so that all ferromagnetic materials become paramagnetic at a sufficiently high temperature. The transition temperature at which all the spin alignment is lost is called the *Curie temperature*. Just below the Curie temperature, the susceptibility can be 10^2 to 10^3 , increasing with decrease in temperature, as the alignment becomes more complete. The Curie temperature is a function of the magnitude of the exchange energy. Cobalt that has the highest exchange energy has the highest Curie temperature, 1400 K. Gadolinium with a small exchange energy has a Curie temperature below room temperature.

Example 16.2 Which of the two solids, cobalt and gadolinium, has the higher saturation magnetization at (i) 0 K, and (ii) 300 K?

Solution (i) At 0 K, the thermal energy kT , which tends to randomize the spins, is zero. Therefore, all the spins remain aligned in both Co and Gd. As the net magnetic moment of Gd is 7 per atom, as compared to 1.7 per atom in the Co crystal, Gd will have the higher saturation magnetization. The actual values will depend on the number of atoms per unit volume in the two cases.

(ii) At 300 K, Gd is above its Curie temperature of 289 K. Hence, Gd will be paramagnetic at 300 K and will have negligible magnetization as compared to Co, which has a much higher Curie temperature.

The interatomic distance can be changed within limits by alloying. For example, manganese in the elemental form is not ferromagnetic. On alloying, the exchange energy can become positive and the spins on neighbouring atoms can become aligned if the Mn–Mn distance is increased by the right amount, so that the ratio of the atomic diameter to the $3d$ orbital diameter falls in the range 1.5–2.0. This happens in *Heusler alloys* Cu_2MnSn and Cu_2MnAl , which are ferromagnetic.

In some compounds, the constituent atoms may be antiferromagnetically coupled, but can have different magnetic moments. This would give rise to a net magnetic moment in each coupling and the sum of the moments of all the couplings can result in magnetization that is comparable in order of magnitude to ferromagnetism. This phenomenon is called *ferrimagnetism* and is compared with ferromagnetism and antiferromagnetism in Fig. 16.2.

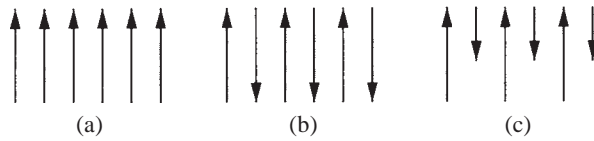


Fig. 16.2 (a) Ferromagnetic, (b) antiferromagnetic, and (c) ferrimagnetic coupling of electron spins in atoms denoted by arrows. The length of an arrow is a measure of the magnetic moment of the atom.

Ferrites of the general formula $\text{Me}^{2+}\text{Fe}_2^{3+}\text{O}_4^{2-}$ can be ferrimagnetic. The divalent metal cation (Me^{2+}) in the formula is usually Fe^{2+} , Ni^{2+} , Zn^{2+} , Mg^{2+} , Co^{2+} , Ba^{2+} , Mn^{2+} or some combination of these ions. These compounds crystallize mostly as the *inverse spinel structure* described in Chap. 5.

As an example of the incomplete cancellation of the magnetic moments, consider the crystal of magnetite, Fe_3O_4 . (Magnetite is the lodestone of the ancient mariners.) The ferric ion, with two $4s$ electrons and one $3d$ electron removed from the neutral iron atom, has a magnetic moment of 5 units and the ferrous ion has 4 units. All the ferric ions in magnetite are antiferromagnetically coupled, so that the net magnetic moment of the compound arises from the ferrous ions only. The magnetization M calculated on the basis of 4 net magnetic moments per Fe_3O_4 is in good agreement with the experimental value for magnetite.

Example 16.3 In nickel ferrite, NiFe_2O_4 , the ferric ions are antiferromagnetically coupled. The magnetization is due to the nickel ions. When zinc is added to nickel ferrite, the magnetization of the crystal increases, even though the zinc ions are not ferromagnetic. Explain how this could happen.

Solution Referring to Chap. 5, the inverse spinel structure of nickel ferrite has all the Ni^{2+} and half of the Fe^{3+} ions in the octahedral sites. The other half of the Fe^{3+} ions are in the tetrahedral sites. The ferric ions in the two sites are antiferromagnetically coupled. When zinc substitutes for the ferric ions in the tetrahedral sites, the antiferromagnetic coupling will not exactly cancel, as the number of ferric ions in the octahedral sites will be more. This would increase the magnetization, as there is an extra contribution from the uncoupled ferric ions of the octahedral sites.

16.4 The Domain Structure

Iron has a high Curie temperature of 1041 K (768°C) and, consequently, almost all the spins remain aligned at room temperature. Yet, at room temperature, an ordinary piece of iron is not magnetic in the absence of an applied field. To explain this discrepancy, Weiss conceived the idea of *magnetic domains*. The

unmagnetized iron crystal consists of a number of domains and within each domain, the spins of all the atoms are aligned. However, the spins of adjacent domains are not parallel. Due to the random orientation of the domains with respect to one another, the magnetic moments cancel out and the crystal as a whole is not magnetic. In the presence of an external field, the domains tend to align themselves with the field, resulting in a large net magnetization of the solid. The process of *magnetization* of the iron crystal is not the aligning of the electron spins of the different atoms, but *it is the alignment of the various domains*, each of which is already ferromagnetic. The domain concept neatly explains the magnetization behaviour of iron. Ample experimental evidence is now available for the presence of magnetic domains.

The *magnetostatic energy* of a ferromagnetic solid can be reduced if a number of domains are arranged such that no poles exist at the surface and no lines of force go out of the material, as illustrated in Fig. 16.3. Hence, a single domain tends to break up into several domains, even if this means an increase in the domain boundary energy. The domain size does not decrease indefinitely, as at some stage the increasing domain boundary energy would oppose this.

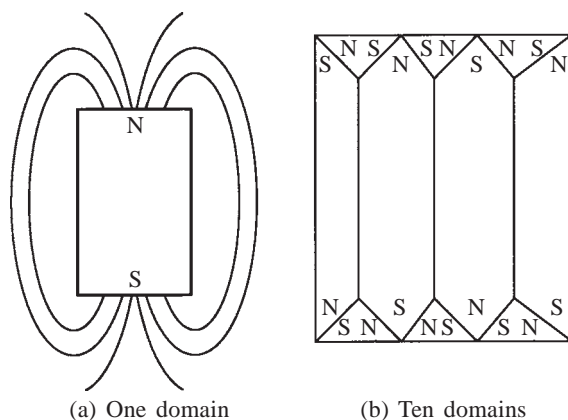


Fig. 16.3 (a) With one domain, the magnetic lines of force go out of the material. (b) The magnetic lines do not go out if a number of domains are present.

16.5 The Hysteresis Loop

The B - H curve for a typical ferromagnetic material is shown in Fig. 16.4. As the applied field H is increased, the magnetic induction B increases slowly at first and then more rapidly. The rate of magnetic induction slows down again, eventually attaining a *saturation value* B_s . With further increase in the magnetic field, there is no increase in the induction. If the field is reversed, the induction decreases slowly at first and reaches a residual value B_r at zero field. B_r represents the amount of *residual induction* left in the specimen after the removal of the field. If the application of the field is continued in the opposite direction, the domains tend to reverse their alignment, so that the remaining

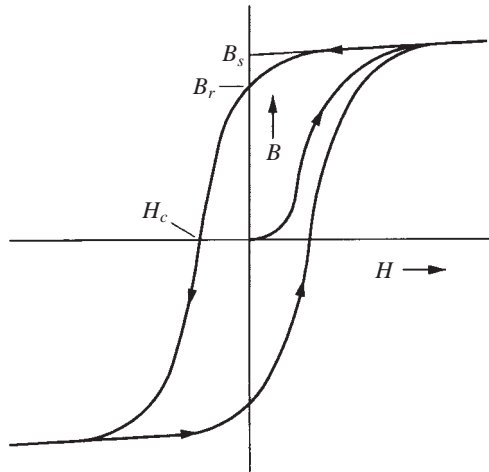


Fig. 16.4 Magnetic induction B as a function of the applied field H for a ferromagnetic material, tracing a hysteresis loop.

induction is lost at a certain value of the reverse field called the *coercive field* H_c . The process of reversal of domains continues to give a net magnetization in the opposite direction. After saturation occurs in this direction, restoring the original field direction completes the *hysteresis loop*.

There are two possible ways to align a random domain structure by applying an external field. One is to rotate a domain in the direction of the field and the other is to allow the growth of the more favourably oriented domains at the expense of the less favourably oriented ones. If we compare the domain structure with the grain structure of a polycrystalline material, the boundaries separating the domains called *domain walls* are the analogue of grain boundaries. The domain boundary energy is about 0.002 J m^{-2} . The domain walls, however, are some two orders of magnitude thicker than the grain boundaries, because there is a gradual transition from one domain orientation to the next across the wall. Also, the domain boundaries can exist within the grain. Analogous to grain growth (Chap. 6), the domain walls can move such that the more favourably oriented domains *grow* at the expense of others. In the earlier stages of magnetization below the saturation region of the hysteresis curve, domain growth is dominant. The growth is more or less complete as the saturation region is approached. Thereafter, the most favourably oriented, fully grown domain tends to *rotate* so as to be in complete alignment with the field direction. The energy required to rotate an entire domain is more than that required to move the domain walls during growth. Consequently, the slope of the B - H curve decreases on approaching saturation.

Each time the hysteresis loop is traversed, *energy equal to the area of the loop is dissipated as heat*. The power loss due to hysteresis in a transformer core is dependent on the number of times the full loop is traversed per second.

Example 16.4 A transformer core is wound with a coil carrying an alternating current at a frequency of 50 Hz. Assuming the magnetization to be uniform throughout the core volume of 0.01 m^3 , calculate the hysteresis loss. The hysteresis loop has an area of 60 000 units, when the axes are drawn in units of $10^{-4} \text{ Wb m}^{-2}$ and 10^2 A m^{-1} .

Solution

Energy lost during each cycle

$$\begin{aligned}
 &= \text{Area of loop} \\
 &= 10^{-4} \times 10^2 \times 60\,000 \\
 &= 600 \text{ J m}^{-3}
 \end{aligned}$$

Volume of the transformer core = 0.01 m^3

Energy lost in the core in each cycle = $600 \times 0.01 = 6 \text{ J}$

Power loss due to hysteresis = $6 \text{ J} \times 50 \text{ Hz} = 300 \text{ W}$

16.6 Soft Magnetic Materials

Soft magnetic materials are used in applications requiring frequent reversals of the direction of magnetization. In soft magnetic materials, the *hysteresis losses must be kept down to a minimum*. When the induction is large for a small applied field, the loop area is small and the hysteresis loss is reduced. The key factor in the design of a soft magnet is then to have *easily moving domain walls*. Analogous to grain boundaries, the domain walls tend to get pinned down by dislocation tangles, impurity atoms, voids and nonmagnetic precipitates and inclusions. A cold-worked material has a high dislocation density. It should be properly annealed to reduce the dislocation density and thereby facilitate easier motion of the domain walls. Soft magnetic materials should be free of impurities and inclusions. Nonmetallic soft magnets such as ferrites and garnets are prone to inherit voids during the process of manufacture by powder compacting. The microstructure is therefore critical in sophisticated applications using ferrites and garnets.

Usually, there are easy and hard magnetization directions in a crystal. As illustrated in Fig. 16.5, iron magnetizes more easily along the $[100]$ direction than along $[111]$, which is the hard direction for iron. This property can be exploited to reduce the area under the hysteresis loop by manufacturing materials with a *preferred orientation of grains*. The preferred orientation can be achieved by a suitable rolling schedule and final recrystallization to produce what is called a *texture*. The same can also be achieved in some cases by casting the liquid alloy in a metal mould. After the initial chilling against the mould walls, long columnar grains form and grow in a direction perpendicular to the mould wall, recall Sec. 9.6. The $\langle 100 \rangle$ directions often being the favoured

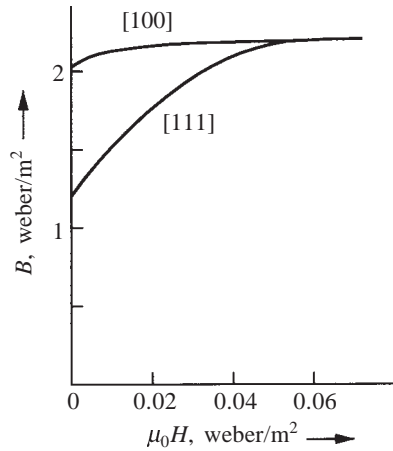


Fig. 16.5 An iron crystal magnetizes more easily along the [100] direction than along [111].

directions of growth, the columnar crystals are aligned along one of the $\langle 100 \rangle$ directions. In the manufacture of ceramic magnets by powder compacting, a field is superimposed across the die cavity before pressing so that the crystalline particles of the powder orient themselves along directions of easy magnetization.

The other source of energy loss in soft magnets is the *eddy current loss*. When the magnetic flux in a medium is changing, an emf is induced. As given by Lenz's law, the induced emf is proportional to the rate of change of flux and hence to the frequency of the alternating current. The induced emf sets up eddy currents in the medium and the power loss due to the eddy currents is equal to V^2/R , where V is the induced emf and R is the resistance of the medium.

Example 16.5 In a 440 V, 50 Hz transformer, the total iron loss is 2300 W. When the applied voltage is 220 V at 25 Hz, the total iron loss is 750 W. Calculate the eddy current loss at the normal voltage and frequency.

Solution The hysteresis loss and the induced emf are proportional to the frequency. The eddy current loss is proportional to the square of the induced emf. Let W_e and W_h be the eddy current and the hysteresis losses at 25 Hz. Then,

$$W_e + W_h = 750 \text{ W}$$

At 440 V and 50 Hz, we have

$$4W_e + 2W_h = 2300 \text{ W}$$

or

$$2W_e = 800 \text{ W}$$

So,

Eddy current loss at the normal voltage and frequency

$$= 4W_e = 1600 \text{ W}$$

Eddy current losses can be minimized by *increasing the resistivity of the magnetic medium*. Iron which used to be the material for transformer cores is now almost entirely replaced by an Fe–Si solid solution with about 4% silicon, which has a substantially higher resistivity than pure iron, as illustrated in Fig. 16.6. The reason for stopping at a concentration of 4% is due to fabrication problems, as higher silicon content tends to make the iron brittle, see the variation in the ductile–brittle transition temperature with silicon content in Fig. 16.6. In addition, the transformer core is laminated such that the resistance of the laminations is much more than that of a one-piece core.

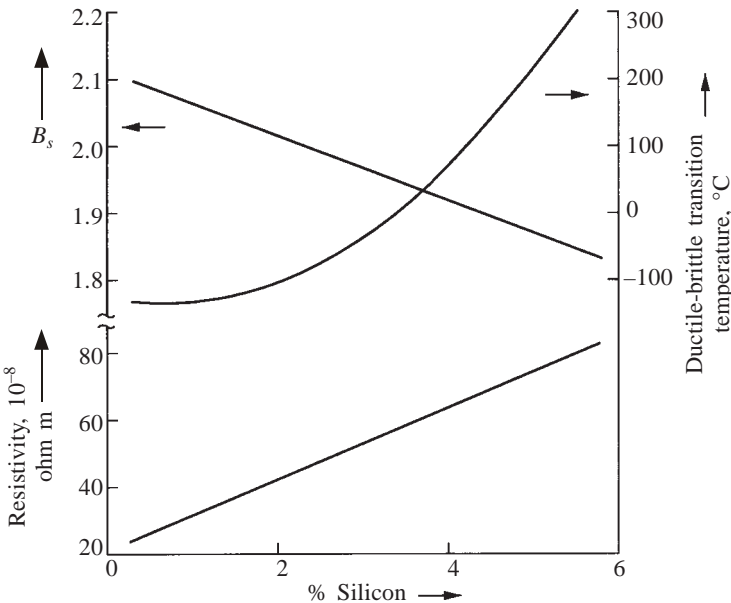


Fig. 16.6 Variation in magnetization, resistivity and ductile–brittle transition temperature as a function of silicon content in Fe–Si alloys.

In recent years, *metallic glasses* produced from iron–base alloys containing 15–25% of (Si + B + C) offer substantial reduction in core losses. Such an alloy, cooled at a rate of $\sim 10^4 \text{ }^{\circ}\text{C s}^{-1}$ from the molten state, does not crystallize but solidifies into a (ribbon-shaped) metallic glass. Owing to the larger concentrations of the impurity atoms, the electrical resistivity is higher than that for the Fe–4% Si alloy, thereby reducing eddy current losses. The absence of grain boundaries in the glassy matrix reduces hysteresis losses. The total iron losses can be reduced to 30–10% of that for the conventional Fe–Si alloy. Such a reduction can save nearly a billion dollars in distribution transformers alone in a developed country like the U.S. Soft magnets made of metallic glass are also used in phonograph cartridges and audio and computer tape heads.

Fe–Si alloys are suitable for operation at power frequencies of 50–60 Hz. They are not suitable in communications equipment, where high sensitivity and

fidelity are required. Fe–Ni alloys such as *Permalloy* and *Supermalloy* are used for this purpose. These alloys have a *high initial permeability* which reduces considerably the area under the hysteresis loop. Hence, these alloys are suitable for higher frequencies. For frequencies exceeding MHz, metals and alloys are generally not suitable as soft magnets, as the eddy current losses are very high. However, thin films of a Permalloy of about 2000 Å in thickness vapour-deposited on a glass substrate have been used. Here, eddy current losses are not significant because of the small thickness. The reversal of magnetization, which occurs by domain rotation, is faster and switching speeds of $\sim 0.1 \mu\text{s}$ can be attained. The ferrimagnetic oxides, *ferrites and garnets*, are very suitable in the high frequency range. Being electrical insulators, they have a much higher resistivity than alloys. This reduces the eddy current losses to a negligible value. The choice depends on the application. Nickel–zinc ferrites are used for *audio and TV transformers*. Magnesium–manganese ferrites with a high resistivity are used as microwave isolators and gyrators in the kHz and MHz range. With a higher manganese to magnesium ratio than in the above, the ferrite has a nearly *rectangular hysteresis loop* and is used for *memory cores in computers*. Garnets such as $\text{Y}_3\text{Fe}_5\text{O}_{12}$ (yttrium–iron–garnet) have a narrow resonance line width and are widely used as *microwave isolators* in the very high frequency GHz range. Table 16.1 lists the properties of some soft magnetic materials.

TABLE 16.1
Properties of Soft Magnetic Materials

Material	Initial relative permeability	Electrical resistivity, ohm m	Hysteresis loss, J m^{-3}	Saturation induction, Wb m^{-2}
Commercial iron	250	0.1×10^{-6}	500	2.2
Fe–4% Si	500	0.6×10^{-6}	100	2.0
Fe–Si oriented	1500	0.6×10^{-6}	90	2.0
Permalloy (45% Ni)	2700	0.55×10^{-6}	120	1.6
Supermalloy (79% Ni, 5% Mo)	100,000	0.65×10^{-6}	21	0.8
Ni–Zn ferrite	200–1000	$\sim 10^6$	35	0.4
Mn–Zn ferrite	2000	$\sim 10^6$	40	0.3

16.7 Hard Magnetic Materials

Hard magnetic materials are used to produce permanent magnets. Hysteresis losses are of no significance here, as no repeated reversals of the magnetization is involved in a permanent magnet. The permanent magnets must have a high residual induction B_r and a large coercive force H_c . The area of the hysteresis loop between B_r and H_c represents the energy required to demagnetize a

permanent magnet. The maximum value of this area is $B_r H_c$, called the *energy product*. It must be as large as possible for permanent magnets. High carbon steels and other low-alloy tungsten and chromium steels are used in the *martensitic condition* as permanent magnets. The same factors that improve the mechanical hardness impart better resistance to domain wall motion in permanent magnets. Among the metallic alloys, the best knowns are the Al–Ni–Co (alnico) alloys. These alloys are directionally solidified and subsequently given a heat treatment in a magnetic field that enhances the desired properties. The heat treatment at 800°C is known to result in a phase separation into two phases of different compositions with different amounts of magnetization. The phase with a larger magnetization precipitates as *very fine elongated particles* embedded in the matrix of the other phase. It is difficult to alter the direction of the magnetic moments in these fine elongated particles, with the result that the energy product $B_r H_c$ is increased. A similar idea is used in the production of fine particle permanent magnets known as ESD (elongated single domain) magnets. The particle thickness here is smaller than the domain wall thickness so that each one of the particles is a single domain. A large field is required to rotate these domains during reversal, as they cannot reverse by the domain growth process. These particles are embedded in a nonmagnetic resin matrix such that they are separated from one another and remain single domains. Among the nonmetallic oxides used as permanent magnets is barium ferrite ($\text{BaO} \cdot 6\text{Fe}_2\text{O}_3$), which has a high coercivity, more than 100 kA m^{-1} , as illustrated in Table 16.2. Recently, permanent magnets of cobalt alloyed with rare earth elements have been developed with a very large energy product in the range of 200 kJ m^{-3} , refer Table 16.2. $\text{Nd}_2\text{Fe}_{14}\text{B}$, an alloy discovered in 1984, has a still higher energy product of 400 kJ m^{-3} . The alloy is cooled from the molten state at a cooling rate of $10^4 \text{ }^\circ\text{C s}^{-1}$ to produce an ultrafine grain size, which is in the same range as the domain size ($\sim 1 \text{ }\mu\text{m}$). The domain walls here are effectively pinned down by the grain boundaries.

TABLE 16.2
Properties of Hard Magnetic Materials

Material	B_r , Wb m^{-2}	H_c , kA m^{-1}	$B_r H_c$, kJ m^{-3}
High carbon steel (martensitic)	0.90	3.98	3.58
Tungsten steel (5% W)	1.05	5.57	5.85
Chromium steel (4% Cr)	0.95	5.17	4.91
Cobalt steel (36% Co)	0.95	18.31	17.40
Al–Ni–Co alloys	0.8–1.2	60–120	48–144
Barium ferrite	0.21	140	29.4
Cobalt rare-earths Co_5 (Sm, Pr)	1.0	200	200

SUMMARY

1. Diamagnetic materials repel magnetic lines of force, while paramagnetic materials attract them. Ferromagnetic materials attract the lines of force very strongly.
2. The permanent magnetic moment of an atom arises from various sources, the most important of which is the spin of electrons in partially filled electron shells. In units of Bohr magneton, the magnetic moment is equal to the number of unpaired electrons. Due to overlapping of energy bands, the net magnetic moment of an atom is less in the solid, as compared to the free state of the atom.
3. The spins of electrons of neighbouring atoms are permanently aligned in ferromagnetic materials, due to bonding forces. Among the common elements, Fe, Ni and Co are ferromagnetic.
4. Ferrimagnetic materials have an antiparallel arrangement of electron spins, but due to incomplete cancellation, there is a net magnetic moment that is comparable in order of magnitude to that of ferromagnetic materials.
5. Even though the spins within a magnetic domain are aligned permanently below the Curie temperature in a ferromagnetic material, the material is magnetic only when an external field aligns the domains.
6. The B - H hysteresis curve can be understood in terms of domain growth and domain rotation.
7. Soft magnetic materials should have low hysteresis losses and low eddy current losses. Easy domain wall motion is the key factor in keeping the hysteresis losses to a minimum. Increasing the electrical resistivity of the magnetic medium reduces eddy current losses.
8. Hard magnetic materials must retain a large part of their magnetization on removal of the applied field. Obstacles to domain wall motion should be provided in permanent magnets, so that the energy product $B_r H_c$ is large.

PROBLEMS

16.1 Write down the equivalence of (i) weber, and (ii) henry in SI base units.

16.2 Show that

$$\frac{\mu}{\mu_0} = 1 + \chi$$

16.3 Show that, for a perfectly diamagnetic material, $M = -H$, $\chi = -1$ and $B = 0$.

- 16.4** Explain why copper crystal is diamagnetic, even though the copper atom has an outer electron configuration of $3d^{10}4s^1$. The diamagnetic susceptibility of copper is -0.5×10^{-5} . For an applied field of 100 kA m^{-1} , find the induction B and the magnetization M .

Answer: 0.126 Wb m^{-2} ; -0.5 A m^{-1} .

- 16.5** Deduce the net magnetic moment of Fe, Co and Ni from the electronic configuration of the atoms.

- 16.6** The Curie temperatures of some elements are: Fe (768°C), Co (1127°C), Gd (16°C), Ni (358°C), and Dy (-168°C). Arrange these in order of increasing exchange energy.

- 16.7** Does a large magnetic moment per atom mean a high Curie temperature? Explain.

- 16.8** Deduce the magnetic moment per formula of the following ferrites: Fe_3O_4 , NiFe_2O_4 , CoFe_2O_4 , and MnFe_2O_4 . In Fe_3O_4 , the ferric ions are antiferromagnetically coupled. All the divalent cations have lost their 4s electrons. Compare the deduced values with the listed values and explain any discrepancy.

Answer: Deduced: 4, 2, 3, and 0 (MnO is antiferromagnetic).

Listed: 4.1, 2.4, 3.7 and 5.

- 16.9** Calculate the saturation magnetization of magnetite. The unit cell of magnetite is cubic with a lattice parameter of 8.37 \AA . There are 16 ferric ions and 8 ferrous ions in the unit cell.

Answer: 510 kA m^{-1} .

- 16.10** The speed of storing and reading out information from a computer core is less than a microsecond. Why is it necessary to use a ferrite for this application rather than a ferromagnetic alloy?

- 16.11** Explain how a high initial permeability in Fe–Ni alloys helps to reduce the area under the hysteresis loop.

- 16.12** Explain the principle underlying the lamination of a transformer core, indicating how you would design it with reference to the direction of the eddy currents.

- 16.13** The total iron loss in a transformer core at normal flux density was measured at 25 Hz and at 50 Hz and was found to be 250 W and 800 W respectively. Calculate (i) the eddy current loss and (ii) the hysteresis loss at 50 Hz.

Answer: (i) 600 W, and (ii) 200 W.

- 16.14** The electrical resistivity of iron increases by a factor of 6, on alloying with 4% silicon. Estimate the decrease in eddy current losses for this change in composition.

Answer: 83%.

MULTIPLE CHOICE QUESTIONS

1. The units of magnetic permeability are
 A. A m^{-1} B. Wb m^{-2} C. H m^{-1} D. $\text{Wb A}^{-1} \text{m}^{-1}$
2. In base SI units, tesla is expressed as
 A. $\text{N A}^{-1} \text{m}^{-1}$ B. $\text{N A}^{-1} \text{m}^{-3}$ C. $\text{kg m}^{-2} \text{s}^{-2} \text{A}^{-1}$ D. $\text{kg s}^{-2} \text{A}^{-1}$
3. In a paramagnetic material of susceptibility equal to 10^{-3} , the induction in Wb m^{-2} at an applied field of 100 kA m^{-1} is ($\mu_0 = 1.257 \times 10^{-6} \text{ H m}^{-1}$)
 A. 0.1257 B. 0.1258 C. $4\pi \times 10^{-7}$ D. none of these
4. The net magnetic moment of Fe atom in BCC crystal ($a = 2.857 \text{ \AA}$) is $2.2\mu_B$ ($\mu_B = 9.273 \times 10^{-24} \text{ A m}^2$). The saturation magnetization of Fe at 0 K is
 A. 0 B. 1750 kA m^{-1} C. 1750 A m^{-1} D. 1750 A m^{-2}
5. Gadolinium has a higher saturation magnetization than Co at
 A. -273°C B. 25°C C. 290°C D. 769°C
6. The Curie temperature of cobalt is
 A. 2000 K B. 1400 K C. 1040 K D. 650 K
7. The total iron loss in a transformer core at 25 Hz and 50 Hz is 250 W and 800 W, respectively. The eddy current loss at 25 Hz is
 A. 100 W B. 150 W C. 200 W D. 600 W
8. Using a permalloy core ($\mu_r = 2700$), a 19 m long 300 turn coil of a conductor giving an induction of 7.5 T is to be made. The current in the coil should be
 A. 0 A B. 52 mA C. 7.4 A D. 140 A
9. The ratio of the atomic radius/ $3d$ shell radius of Mn, which normally exhibits antiferromagnetic behaviour, is
 A. 1.47 B. 1.63 C. 1.82 D. 1.98
10. The temperature of the antiferromagnetic-to-paramagnetic transition is called
 A. antiferromagnetic Curie temp. B. Curie–Weiss temp.
 C. Neel temp. D. Debye temp.
11. The factors that obstruct domain wall motion in Fe are
 A. dislocation tangles B. impurity atoms
 C. voids D. nonmagnetic inclusions
12. The transition from the ferromagnetic to the paramagnetic state is named after
 A. Curie B. Curie–Weiss C. Neel D. Debye

- | | | | | |
|----------------|-------|-------|-------|-------|
| 1. C, D | 2. D | 3. B | 4. B | 5. A |
| 6. B | 7. B | 8. D | 9. A | 10. C |
| 11. A, B, C, D | 12. A | 13. A | 14. D | 15. A |
| 16. B | 17. A | 18. B | 19. C | 20. C |
| 21. A | | | | |

Sources for Experimental Data

F.N. Bradley, *Materials for Magnetic Functions*, Hayden, New York (1971).

ASM International, *Metals Handbook*, 10th ed, Vol. 2: *Special-Purpose Materials*, Materials Park, Ohio (1990), pp. 761–803.

Suggestions for Further Reading

F. Brailsford, *An Introduction to the Magnetic Properties of Materials*, Longmans, London (1968).

L. Solymar and D. Walsh, *Lectures on the Electrical Properties of Materials*, Oxford University Press, Oxford (1984).

When classifying solids on the basis of their band structure, we referred to the group of solids, called insulators, which have an energy gap of 3 eV or more, recall Sec. 15.1. The large magnitude of the energy gap in an ideal insulator precludes the possibility of electrons being excited from the valence band to the conduction band by thermal means, much less so by an externally applied electric field. Insulators are therefore very poor conductors of electricity, recall the resistivity range for insulators from Table 14.1. Insulators are known as dielectrics. Dielectric materials find extensive use in the electrical industry for insulation purposes and as capacitors.

In this chapter, polarization processes in dielectric materials and their temperature and frequency dependence are discussed. Besides, the causes of electric breakdown are described. In the end, the nature of ferroelectric crystals is considered briefly.

Units

Quantity	SI units	
	<i>Unit</i>	<i>Symbol</i>
Flux density D	coulomb per square metre	C m^{-2}
Electric field strength E	volt per metre	V m^{-1}
Dielectric strength	or	or
	newton per coulomb	N C^{-1}
Capacitance C	farad or coulomb per volt	F or C V^{-1}
Dielectric constant of free space ϵ_0	farad per metre	F m^{-1}
Relative dielectric constant ϵ_r	—	—
Dipole moment	coulomb metre	C m
Polarization P	coulomb per square metre	C m^{-2}
Saturation polarization P_s		
Electronic polarizability	farad metre squared	F m^2
Frequency ν	hertz (per second)	$\text{Hz (s}^{-1}\text{)}$

Constants

Dielectric constant of free space ϵ_0	$= 8.854 \times 10^{-12} \text{ F m}^{-1}$
Electronic charge e	$= 1.602 \times 10^{-19} \text{ C}$

17.1 Polarization

The relation between the electric flux density D (charge per unit area in units of C m^{-2}) at a point in a material and the electric field strength E (force per unit charge in units of N C^{-1} or V m^{-1}) at that point in space is given by

$$D = \epsilon_r \epsilon_0 E \quad (17.1)$$

where ϵ_0 is the dielectric constant or permittivity of free space (vacuum) and ϵ_r is the relative dielectric constant or relative permittivity of the material. ϵ_r is dimensionless and is a property of the material related to its atomic structure. Its value is 1 for free space and is greater than one for all materials. In SI units, $\epsilon_0 = 8.854 \times 10^{-12}$ farad per metre. ϵ_0 has this particular value as a result of conversion from cgs units to SI units.

$$\epsilon_0 = \frac{D}{\epsilon_r E} = \frac{\text{C m}^{-2}}{\text{V m}^{-1}} = \text{F m}^{-1}$$

When an electric field is applied to a solid containing positive and negative charges, the positive charges are displaced in the direction of the field towards the negative end, while the negative charges are displaced in the opposite direction. This displacement produces local dipoles throughout the solid. *The dipole moment per unit volume of the solid is the sum of all the individual dipole moments within that volume and is called the polarization P of the solid.* As the polarization measures the additional flux density arising from the presence of the material as compared to free space, it has the same units as D and is related to it as follows:

$$D = \epsilon_0 E + P \quad (17.2)$$

As may be seen by comparing Eq. (17.2) with Eq. (16.3), polarization P is the electrical analog of magnetization M . Combining Eqs. (17.1) and (17.2), we can write

$$P = \epsilon_0 (\epsilon_r - 1) E \quad (17.3)$$

Macroscopically, the polarization or the relative dielectric constant can be measured, using a parallel-plate capacitor. When a voltage V is applied to the capacitor, with the plates separated by vacuum, a charge develops on the plates. The capacitance C of the capacitor is a measure of this charge and is defined by

$$C = \frac{\epsilon_0 A}{d} \quad (17.4)$$

where A is the area of the parallel plates and d is the distance of separation between them. If a dielectric material is inserted between the plates, the charge on the plates increases due to polarization in the material. The capacitance is now given by

$$C = \epsilon_r \epsilon_0 \frac{A}{d} \tag{17.5}$$

ϵ_r and hence P can be determined by measuring the capacitance with and without the dielectric.

Example 17.1 Calculate the relative dielectric constant of a barium titanate crystal, which, when inserted in a parallel plate condenser of area $10 \text{ mm} \times 10 \text{ mm}$ and distance of separation of 2 mm , gives a capacitance of 10^{-9} F .

Solution Substituting the given values in Eq. (17.5), we obtain

$$\begin{aligned} \epsilon_r &= \frac{10^{-9} \times 2 \times 10^{-3}}{8.854 \times 10^{-12} \times 10^2 \times 10^{-6}} \\ &= 2259 \end{aligned}$$

Polarization occurs due to several atomic mechanisms. *Electronic polarization* (Fig. 17.1a) is the result of the displacement of the positively charged nucleus and the (negative) electrons of an atom in opposite directions on application of an electric field. On applying a field, the electron cloud around the nucleus readily shifts towards the positive end of the field. Such a shift results in a *dipole moment* within the atom, as a certain distance now separates the nucleus and the centre of the electron cloud. The extent of this shift is proportional to the field strength. As the dipole moment is defined as the product of the charge and the shift distance, it is also proportional to the field strength. The constant of proportionality is called the *electronic polarizability* α_e of the atom. For the inert gases, this polarizability increases with increasing volume of the atom, as illustrated in Table 17.1. Electronic polarizability is independent of temperature. Monoatomic gases exhibit only this kind of polarization.

TABLE 17.1
Electronic Polarizability α_e of Inert Gases

Inert gas	He	Ne	Ar	Kr	Xe
$\alpha_e, 10^{-40} \text{ F m}^2$	0.18	0.35	1.43	2.18	3.54

During chemical bonding, the atoms may acquire an excess negative or positive charge and form an ionic bond. When an electric field is applied to an ionic solid, cations and anions get displaced in opposite directions, see

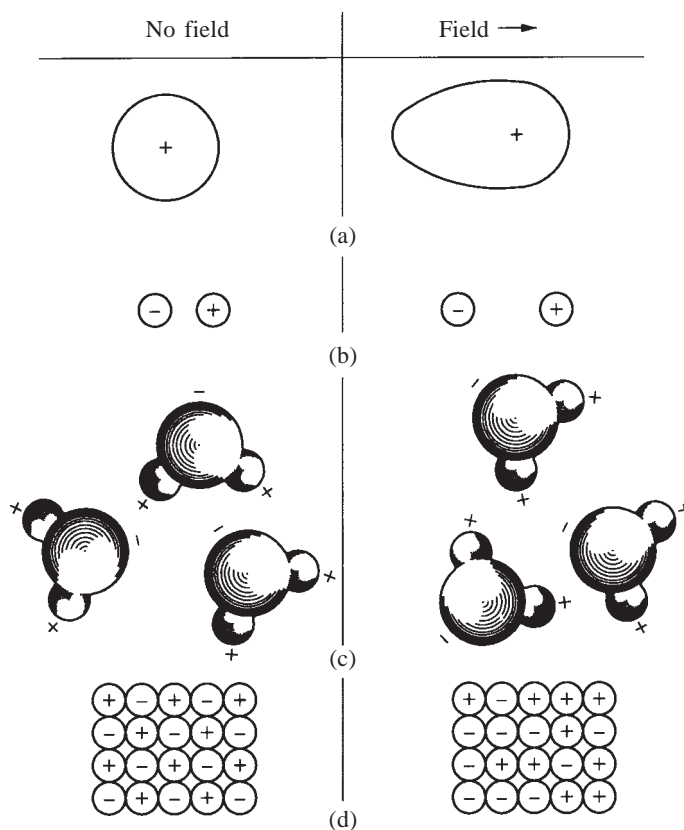


Fig. 17.1 Various polarization processes: (a) electronic polarization, (b) ionic polarization, (c) orientation polarization, and (d) space charge polarization.

Fig. 17.1b. The *ionic polarizability* is due to this shift of the ions relative to other oppositely-charged neighbours. It should be distinguished from electronic polarization, where the electron cloud of an atom shifts with reference to its own nucleus. Ionic polarization is also independent of temperature.

In methane molecule (CH_4), the centre of the negative and the positive charges coincide, so that there is no permanent dipole moment. On the other hand, in CH_3Cl , the positive and the negative charges do not coincide. The electronegativity of chlorine being more than that of hydrogen (recall Table 4.4), the chlorine atom pulls the bonding electrons to itself more strongly than hydrogen. So, this molecule carries a dipole moment even *in the absence of an electric field*. When an electric field is applied on such molecules, they tend to align themselves in the applied field, see Fig. 17.1c. (Recall that atoms with permanent magnetic moments tend to align themselves with the applied magnetic field giving rise to paramagnetism.) The polarization due to this alignment is called *orientation polarization* and is dependent on temperature. With increasing temperature, the thermal energy tends to randomize the alignment.

The fourth type of polarization is called the *space charge polarization* and occurs due to the accumulation of charges at the electrodes or at the interfaces in a multiphase material, see Fig. 17.1d. The ions diffuse over appreciable distances in response to the applied field, giving rise to a redistribution of charges in the dielectric medium.

The *total polarization* of a material is the sum of the contributions from the various sources described above:

$$P_{\text{total}} = P_e + P_i + P_o + P_s \quad (17.6)$$

where the subscripts on the right refer to the four types: electronic, ionic, orientation and space-charge polarization.

17.2 Temperature and Frequency Effects

On application of an electric field, a polarization process occurs as a function of time. The polarization $P(t)$ as a function of time t is given by

$$P(t) = P \left[1 - \exp \left(- \frac{t}{t_r} \right) \right] \quad (17.7)$$

where P is the maximum polarization attained on prolonged application of a static field and t_r is the relaxation time for the particular polarization process. Recall Eq. (10.15), where the relaxation arising from anelastic processes was discussed. The relaxation time t_r is a measure of the time scale of a polarization process. It is the time taken for a polarization process to reach 0.63 of the maximum value.

The relaxation times vary widely for different polarization processes. Electronic polarization is *extremely rapid* and is essentially complete at the instant the voltage is applied. Even when the frequency of the applied voltage is very high in the optical range ($\sim 10^{15}$ Hz), electronic polarization occurs during every cycle of the applied voltage.

Ionic polarization is slower than electronic polarization, as the displacement involved here is that of the much heavier ion, as compared to the electron cloud in the above case. The frequency with which ions can be displaced over a small fraction of the interatomic distance will be of the same order as the lattice vibration frequency ($\sim 10^{13}$ Hz). If an electric field of frequency in the optical range ($\sim 10^{15}$ Hz) is now applied, the ions do not respond at all, as the time required by an ion for one vibration is 100 times larger than the period of the applied voltage. So, at optical frequencies, there is no ionic polarization. If the frequency of the applied voltage is less than 10^{13} Hz, the ions have enough time to

respond during each cycle. This is illustrated in Fig. 17.2, where the polarization consists of only the electronic part at optical frequencies and increases to include the ionic part, as the frequency is reduced to 10^{13} Hz in the infrared range.

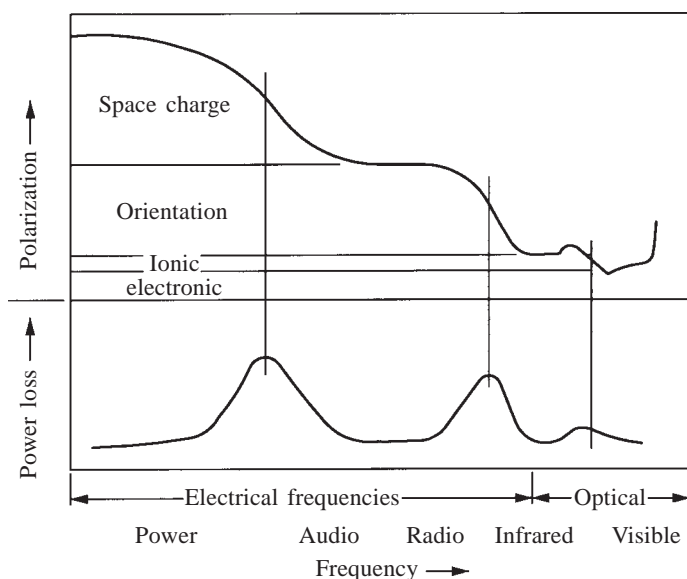


Fig. 17.2 Frequency dependence of polarization processes and peak power losses corresponding to each process.

Orientation polarization is even slower than ionic polarization. The relaxation time for orientation polarization in liquid propyl alcohol is about 10^{-10} s and in solid ice, it is 3×10^{-6} s. It is evidently easier for the polar molecules to reorient themselves in a liquid as compared to a solid. Orientation polarization occurs, when the frequency of the applied voltage is in the audio range, Fig. 17.2.

Space charge polarization is the *slowest process*, as it involves the diffusion of ions over several interatomic distances. The relaxation time for this process is related to the frequency of successful jumps of ions under the influence of the applied field, a typical value being 10^2 Hz. Correspondingly, space charge polarization occurs at machine frequencies (50–60 Hz).

Referring to Fig. 17.2, all the four types of polarization are present at machine frequencies. As the frequency is increased, space charge, orientation and ionic polarization become inoperative in that order. When several polarization processes occur in a material, it follows that the dielectric constant will decrease with increasing frequency of the applied voltage. The dielectric constants of some typical dielectric materials are listed in Table 17.2 at frequencies of 60 Hz and 10^6 Hz.

TABLE 17.2
Properties of Some Dielectric Materials

Material	ϵ_r		$\tan \delta$, 10^6 Hz	Dielectric strength, 10^6 V m ⁻¹
	60 Hz	10^6 Hz		
Electrical porcelain	6	6	0.02	5
Steatite, MgO·SiO ₂	6	6	0.001	12
Fused silica	4	3.8	0.0001	10
Soda-lime-glass	7	7	0.005	10
Mica	8	5	0.0005	100
Nylon 6, 6	4	3.5	0.02	15
Polyethylene	2.3	2.3	0.0004	4
Polyvinylchloride (plasticized)	7	3.4	0.05	2
Vulcanized rubber	4	2.7	0.003	25
Bakelite	4.4	4.4	0.028	15
Transformer oil	5	2.5	0.0001	10

When the period of the applied voltage is much larger than the relaxation time of a polarization process, polarization is essentially complete at any instant during each cycle. The charging current is 90° advanced in relation to the voltage, as illustrated in Fig. 17.3a, so that *no electrical energy is lost during charging*.

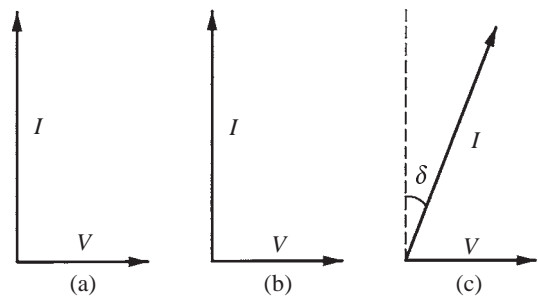


Fig. 17.3 (a) and (b) There is no energy loss when the period of the ac voltage does not match the relaxation time of a process. (c) When there is matching, energy loss occurs, with the current leading the voltage by less than 90° .

When the period of the applied voltage is much shorter than the relaxation time for a polarization process, the polarization does not occur at all. Here again, the charging current is 90° advanced of the applied voltage, see Fig. 17.3b. When the period is in the same range as the relaxation time, *resonance* occurs. Here, the current leads the voltage by $(90 - \delta)$, as illustrated in Fig. 17.3c. δ is called the loss angle and $\tan \delta$ is taken as a measure of the electrical loss due to resonance. The current can be factorized into a component at 90° to the voltage as in an ideal

capacitor and another component parallel to the voltage. The parallel component is the real part and results in I^2R loss. The $\tan \delta$ values of some dielectric materials at a frequency of 10^6 Hz are listed in Table 17.2.

The effect of temperature on the relative dielectric constant of a material can be two-fold. In orientation polarization, the *randomizing action of thermal energy* decreases the tendency for the permanent dipoles to align themselves in the applied field. This results in a decrease in the dielectric constant with increasing temperature. For example, ϵ_r for solid HCl decreases inversely with temperature between 100 K and 160 K from 19 to 14. At the melting point of 160 K, it drops abruptly to 12, as there is an expansion on melting and the number of molecules per unit volume decreases.

The other effect of temperature is to *facilitate the diffusion of ions* in space charge polarization. Thermal energy may also *aid in overcoming the activation barrier* for the orientation of polar molecules in the direction of the field. For example, the relatively large polar molecule nitrobenzene is not able to reorient itself in the solid state under an applied field and, therefore, does not exhibit orientation polarization. However, on melting, the molecules have sufficient thermal energy to orient themselves in the applied field. Correspondingly, the relative dielectric constant increases here from 3 to 37 on melting, which is in contrast to the behaviour of HCl described above.

17.3 Electric Breakdown

In practical applications, the failure or breakdown of a dielectric material is of great concern to the engineer. In a dielectric, the charge displacement increases with increasing field strength. Beyond a critical value of the field strength, there is an electric breakdown due to the physical deterioration of the material. The *dielectric strength* is defined as the breakdown voltage per unit thickness of the material. The dielectric strengths of some common insulating materials are listed in Table 17.2. Mica has one of the highest dielectric strengths and is widely used for insulation purposes.

The *intrinsic breakdown* of a dielectric material is due to the excitation of electrons into the conduction band across the energy gap under conditions of excessive voltage. The excited electrons moving under a high accelerating force can excite more electrons in turn, resulting in an avalanche of conducting electrons and consequent physical breakdown. It is difficult to see how this can occur in an ideal dielectric, where the energy gap is large. However, *impurities in the dielectric can create additional energy levels that lie in the energy gap* and can help in the excitation of electrons into the conduction band. An increase in temperature aids the thermal excitation of electrons and can bring about the intrinsic breakdown.

Thermal breakdown is due to the attainment of an excessive temperature in the dielectric. The energy loss referred to earlier has to be dissipated as heat. If the heat dissipated is less than the heat generated, there is a progressive increase in the temperature of the dielectric, which may melt eventually.

Defect breakdown is due to cracks and pores at the surface. To decrease the possibility of surface shorting, insulators are designed with lengthened surface paths. *Moisture* from the atmosphere can collect on the surface discontinuities and result in breakdown. Glazing is done on ceramic insulators to make the surface nonabsorbent. Gases can collect at pores and cracks and the breakdown can occur due to a *gas discharge*. A graphic case of surface breakdown of an insulator is shown in Fig. 17.4.

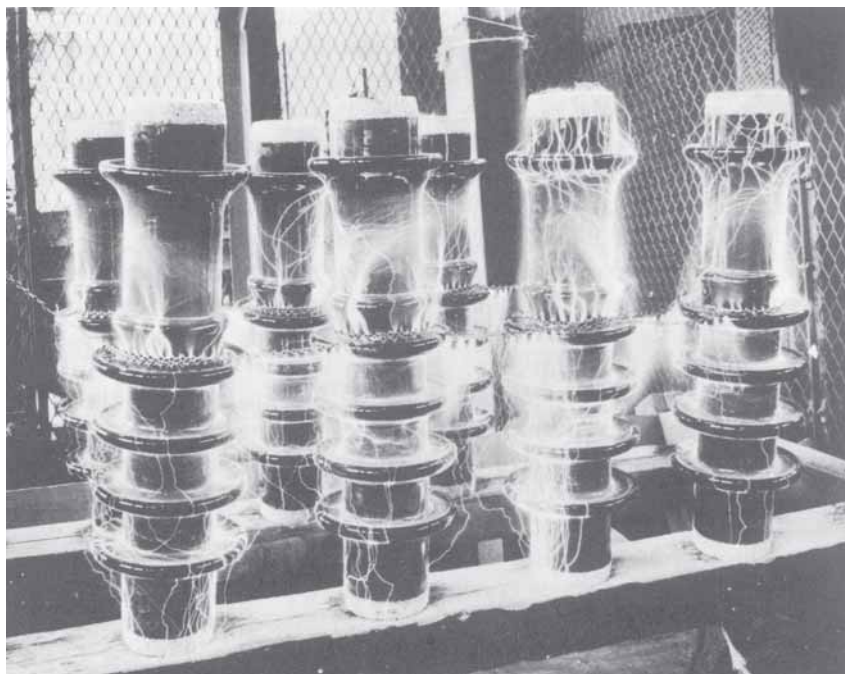


Fig. 17.4 Surface breakdown of an insulator. (L.H. Van Vlack, *Physical Ceramics for Engineers*, by permission from Addison-Wesley, Reading, Mass.)

17.4 Ferroelectric Materials

In materials known as ferroelectrics, the dielectric constants are some two orders of magnitude larger than those in ordinary dielectrics. *Barium titanate* is a ferroelectric with a relative dielectric constant of over 2000, compared to less than 10 for ordinary dielectrics listed in Table 17.2. The difference in the magnetic susceptibility between ferromagnetic and paramagnetic materials bears a direct analogy to this difference in the values of the dielectric constants. Following the nomenclature in magnetism, materials of very large dielectric constants are called *ferroelectrics*. As in the ferromagnetic phenomenon, *the electric dipoles in a ferroelectric solid are all aligned in the same direction, even in the absence of an electric field.*

The ferroelectric phenomenon is discussed here with reference to the classical example of barium titanate, BaTiO_3 . The cubic unit cell of barium titanate crystal is shown in Fig. 17.5a. Barium ions are at the body corners, the

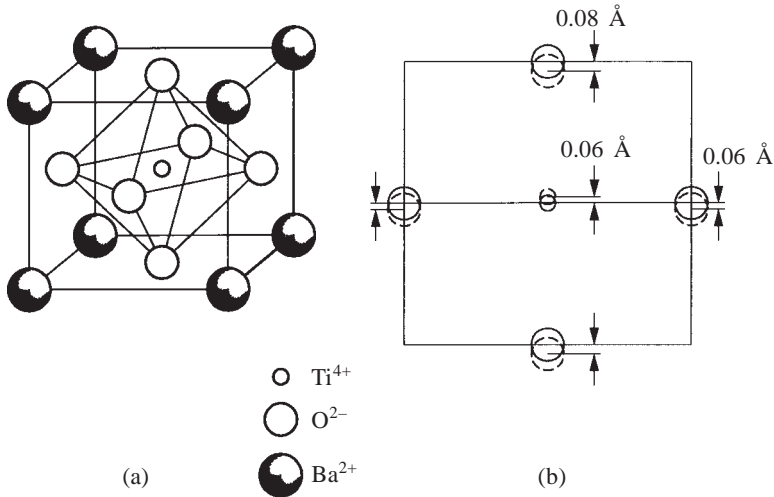


Fig. 17.5 (a) Cubic unit cell of BaTiO_3 crystal. (b) The dashed circle in the middle of the front view shows the shifting of the titanium ion, on cooling through the Curie temperature. The shift of the oxygen anions is also shown.

oxygen anions are at the face centres and the titanium ion is in the octahedral void at the body centre. Only one out of four octahedral voids in the unit cell is occupied and this corresponds to the chemical formula, with one titanium for every four species of the other kinds: one barium plus three oxygen. Above 120°C , barium titanate is a cubic crystal with the ion locations as described above. In this state, the centres of the negative and the positive charges coincide and there is no spontaneous dipole moment. If the crystal is cooled to below the (ferroelectric) Curie temperature of 120°C , the titanium ion shifts to one side of the body centre as shown dotted in the front view of Fig. 17.5b. There is also a displacement of the neighbouring oxygen anions. The crystal *transforms* from a cubic to a tetragonal structure on cooling through 120°C . The c/a ratio of the tetragonal cell is $4.03 \text{ \AA}/3.98 \text{ \AA} = 1.012$. The centres of the positive and the negative charges do not coincide any longer and local dipoles are created throughout the crystal. The dipoles of neighbouring unit cells are all aligned resulting in a large polarization in the solid.

Example 17.2 Calculate the polarization of a BaTiO_3 crystal. The shift of the titanium ion from the body centre is 0.06 \AA . The oxygen anions of the side faces shift by 0.06 \AA , while the oxygen anions of the top and bottom faces shift by 0.08 \AA , all in a direction opposite to that of the titanium ion.

Solution The dipole moments due to the effective number of each type of ion in the unit cell are due to

- (i) two O^{2-} on four side faces: $2 \times 2 \times 1.6 \times 10^{-19} \times 0.06 \times 10^{-10} \text{ C m}$;
- (ii) one O^{2-} on top and bottom: $1 \times 2 \times 1.6 \times 10^{-19} \times 0.08 \times 10^{-10} \text{ C m}$; and
- (iii) one Ti^{4+} at body centre: $1 \times 4 \times 1.6 \times 10^{-19} \times 0.06 \times 10^{-10} \text{ C m}$

$$\text{Total} = 1.02 \times 10^{-29} \text{ C m}$$

Polarization is the sum of the dipole moments per unit volume. Ignoring the effect due to the barium ions, we get

$$P = \frac{1.02 \times 10^{-29}}{4.03 \times 3.98^2 \times 10^{-30}} \text{ C m}^{-2} = 0.16 \text{ C m}^{-2}.$$

At room temperature, a BaTiO_3 crystal ordinarily exhibits no net polarization, in the absence of an external field, even though the dipoles of adjacent unit cells are aligned. The reason for this can be understood by visualizing *ferroelectric domains* in the same way as ferromagnetic domains discussed in Sec. 16.4. The application of the electric field tends to align the domains in the direction of the field and we observe all the phenomena associated with the hysteresis loop such as domain rotation and domain growth. Figure 17.6 illustrates a ferroelectric hysteresis loop. The spontaneous

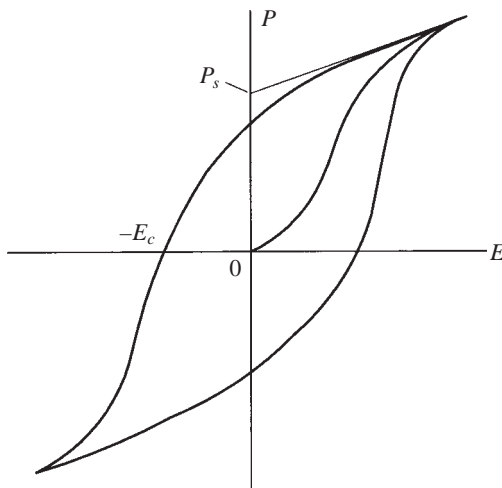


Fig. 17.6 The hysteresis loop for a ferroelectric material, depicting applied field E versus polarization P . P_s is spontaneous polarization.

polarization P_s is obtained by extrapolating the linear region of the curve backwards to zero electric field. The ferroelectric Curie temperatures and the spontaneous polarization P_s below the Curie point of some ferroelectric crystals are listed in Table 17.3.

TABLE 17.3
Properties of Some Ferroelectric Crystals

Crystal	Curie temperature, K	Spontaneous polarization, $P_s, \text{C m}^{-2}$
BaTiO ₃	393	0.26
KNbO ₃	712	0.30
PbTiO ₃	763	0.50
KH ₂ PO ₄	123	0.05

Ferroelectric crystals always exhibit the *piezoelectric property*, which is the mechanical response of a crystal to an electric field or the electrical response to a mechanical stress. With no external field, the centres of the positive and the negative charges are separated by a distance d . If a compressive stress is applied to the crystal, d decreases and a potential difference V develops between the two ends of the crystal. V will be of the opposite sign if a tensile stress is applied. If an external voltage is applied to a ferroelectric crystal, the separation distance d increases (or decreases), thereby elastically straining the crystal. The most important use of BaTiO₃ is as a piezoelectric crystal in applications such as microphones, phonograph pickups, strain gauges and sonar devices. The high dielectric constant of ferroelectric crystals is also useful for storing energy in small-sized capacitors in electrical circuits.

SUMMARY

1. The charge stored in a capacitor increases if a dielectric material is inserted between the capacitor plates. This is due to the polarization in the dielectric.
2. Several polarization processes such as electronic, ionic, orientation and space charge polarization are known. They have widely varying relaxation times.
3. The dielectric constant of a material is dependent on the frequency of the alternating field applied and the temperature.
4. Electrical energy loss is a maximum when the relaxation time of a polarization process matches the period of the alternating field.
5. Impurities, surface flaws such as cracks and pores and absorption of moisture and gases on the surface are the main causes of electric breakdown of a dielectric material.
6. Ferroelectric crystals are in the spontaneously polarized state and have very large dielectric constants. They are useful as piezoelectric crystals and as capacitors.

PROBLEMS

- 17.1** The electric field strength can be expressed as V m^{-1} or as N C^{-1} . Express these in base SI units and show that they are equivalent.
- 17.2** Derive the units of α_e from that of E and the dipole moment.
- 17.3** Find the capacitance of a layer of Al_2O_3 that is $0.5\text{ }\mu\text{m}$ thick and 2000 mm^2 of surface area. $\epsilon_r = 8$.
Answer: $2.83 \times 10^{-7}\text{ F}$.
- 17.4** The electronic polarizability of helium is $0.18 \times 10^{-40}\text{ F m}^2$. Calculate its relative dielectric constant at 0°C and 1 atm pressure.
Answer: 1.000 055.
- 17.5** Estimate the shift of the electron cloud with respect to the nucleus in an argon atom, when a field of 10^5 V m^{-1} is applied.
Answer: $5 \times 10^{-18}\text{ m}$.
- 17.6** The dielectric constant of polyethylene is independent of temperature, while that of polyvinylchloride is not. Explain this difference in behaviour on the basis of their monomer structures.
- 17.7** When ice melts into water, the dielectric constant increases, in contrast to the decrease observed during the melting of HCl . Explain why this is so.
- 17.8** What is the variation expected in the loss factor in a sodium silicate glass, when some of the Na^+ ions are replaced by Rb^+ ions? Explain.
- 17.9** Compare the hysteresis loops in Figs. 16.4 and 17.6. Explain why there is a constant saturation induction B_s in Fig. 16.4, and no saturation value of polarization in Fig. 17.6.

MULTIPLE CHOICE QUESTIONS

- The units of relative dielectric constant is
 A. dimensionless B. F m^{-1} C. C V^{-1} D. F C^{-1}
- The units of ϵ_0 are
 A. C m^{-2} B. H m^{-1} C. F m^{-1} D. dimensionless
- The capacitance of a capacitor with a layer of Al_2O_3 ($\epsilon_r = 8$) of $0.5\text{ }\mu\text{m}$ thick and 1000 mm^2 area is
 A. $0.283\text{ }\mu\text{F}$ B. $0.142\text{ }\mu\text{F}$ C. 17 nF D. 3540 pF
- Expressed in base SI units, the electronic polarizability has the units
 A. F m^2 B. $\text{C V}^{-1}\text{ m}^{-2}$ C. $\text{A}^2\text{ kg}^{-1}\text{ s}^4$ D. $\text{A kg}^{-1}\text{ m}^{-2}\text{ s}^2$

5. If the electronic polarization of W is $4 \times 10^{-7} \text{ C m}^{-2}$, the average displacement of the electrons relative to the nucleus is (W : at.no. 74, BCC, $a = 3.16 \text{ \AA}$)
 A. 3.9 \AA B. $5.3 \times 10^{-19} \text{ m}$ C. $10.6 \times 10^{-19} \text{ m}$ D. $3.9 \times 10^{-17} \text{ m}$
6. The shift of the electron cloud with respect to the nucleus of a helium atom when a field of 10^5 V m^{-1} is applied is (α_e for He = $0.18 \times 10^{-40} \text{ F m}^2$)
 A. $5.6 \times 10^{-18} \text{ m}$ B. $6.2 \times 10^{-19} \text{ m}$ C. $1.7 \times 10^{-10} \text{ m}$ D. $1.7 \times 10^{-18} \text{ m}$
7. Ionic polarization
 A. decreases with temperature
 B. increases with temperature
 C. may increase or decrease with temperature
 D. is independent of temperature
8. With increase in temperature, the orientation polarization in general
 A. increases B. decreases C. is constant D. none of these
9. Among the common dielectric materials, the highest dielectric strength is possessed by
 A. mica B. transformer oil
 C. PVC D. polyethylene
10. During melting, the relative dielectric constant
 A. always increases B. always decreases
 C. may increase or decrease D. none of these
11. The fraction of tetrahedral voids occupied by Ti^{4+} ion in BaTiO_3 crystal is
 A. $1/2$ B. $1/4$ C. $1/8$ D. none of these
12. In the polarization versus field strength plot for a ferroelectric crystal, P_s stands for
 A. space charge polarization B. saturation polarization
 C. spontaneous polarization D. none of these
13. A piezoelectric crystal has a Young's modulus of 130 GPa. The uniaxial tensile stress that must be applied to increase its polarization from 550 to 555 C m^{-2} is
 A. 1.171 GPa B. 1171 MPa C. 2600 MPa D. 1.182 GPa

Answers

- | | | | | |
|-------|-------|-------|------|-------|
| 1. A | 2. C | 3. B | 4. C | 5. B |
| 6. A | 7. D | 8. B | 9. A | 10. C |
| 11. D | 12. C | 13. D | | |

Sources for Experimental Data

H.L. Saums and W.W. Pendleton, *Materials for Electrical Insulating and Dielectric Functions*, Hayden, New York (1971).

R.W. Sillars, *Electrical Insulating Materials and Their Application*, P. Peregrinus Ltd., Stevenage, UK (1973).

Suggestions for Further Reading

A.J. Dekker, *Electrical Engineering Materials*, Prentice-Hall of India, New Delhi (1977), Chaps. 2–3.

W.D. Kingery, H.K. Bowen and D.R. Uhlmann, *Introduction to Ceramics*, Wiley, New York (1976), Chap. 18.

APPENDIX

I Properties of Elements

Element	Symbol	Atomic number	Density, 10^3 kg m^{-3}	Molar volume, 10^{-8} m^3	Thermal expansion, 10^{-6} K^{-1}	Young's modulus, GN m^{-2}	Melting point, $^{\circ}\text{C}$
Actinium	Ac	89	—	—	14.9	35	1050
Aluminium	Al	13	2.70	9.99	23.1	71	660
Americium	Am	95	11.7	20.77	—	—	—
Antimony	Sb	51	6.62	18.4	10.9	55	630
Argon	Ar	18	—	—	—	—	-189
Arsenic	As	33	5.72	13.09	4.28	39	817
Astatine	At	85	—	—	—	—	302
Barium	Ba	56	3.5	39.0	18.8	12.7	714
Berkelium	Bk	97	—	—	—	—	—
Beryllium	Be	4	1.85	4.90	11.5	289	1277
Bismuth	Bi	83	9.80	21.3	13.41	34	271
Boron	B	5	2.34	4.62	8.3	440	2030
Bromine	Br	35	3.12	25.6	—	—	-7
Cadmium	Cd	48	8.65	13.0	30.6	62	321
Calcium	Ca	20	1.55	25.86	22.4	19.5	838
Californium	Cf	98	—	—	—	—	—
Carbon (gr)	C	6	2.25	5.33	3.8	8.3	3550
Cerium	Ce	58	6.77	17.03	8.5	30	804
Cesium	Cs	55	1.90	69.84	97	1.75	28
Chlorine	Cl	17	—	—	—	—	-101
Chromium	Cr	24	7.19	7.23	8.4	243	1875
Cobalt	Co	27	8.85	6.66	12.4	206	1495
Copper	Cu	29	8.96	7.09	16.7	124	1083
Curium	Cm	96	—	—	—	—	—
Dysprosium	Dy	66	8.55	19.01	10.0	63	1407
Einsteinium	Es	99	—	—	—	—	—
Erbium	Er	68	9.15	18.28	12.3	73	1497
Europium	Eu	63	5.25	28.98	33.1	15	826

(Cont.)

APPENDIX I (cont.)

Element	Symbol	Atomic number	Density, 10 ³ kg m ⁻³	Molar volume, 10 ⁻⁸ m ³	Thermal expansion, 10 ⁻⁶ K ⁻¹	Young's modulus, GN m ⁻²	Melting point, °C
Fermium	Fm	100	—	—	—	—	—
Fluorine	F	9	—	—	—	—	-220
Francium	Fr	87	—	—	102	1.7	27
Gadolinium	Gd	64	7.86	20.01	8.28	56	1312
Gallium	Ga	31	5.91	11.8	18.1	92.5	30
Germanium	Ge	32	5.32	13.64	5.75	99	937
Gold	Au	79	19.32	10.20	14.1	78	1063
Hafnium	Hf	72	13.09	13.64	6.01	137	2222
Helium	He	2	—	—	—	—	-270
Holmium	Ho	67	6.79	24.3	10.7	67	1461
Hydrogen	H	1	—	—	—	—	-259
Indium	In	49	7.31	15.71	31.4	10.5	156
Iodine	I	53	4.94	25.7	—	—	114
Iridium	Ir	77	22.5	8.54	6.63	528	2454
Iron	Fe	26	7.87	7.1	11.7	210	1535
Krypton	Kr	36	—	—	—	—	-157
Lanthanum	La	57	6.19	22.44	10.4	38	920
Lawrencium	Lw	103	—	—	—	—	—
Lead	Pb	82	11.36	18.27	29.0	15.7	327
Lithium	Li	3	0.53	12.99	45	11.5	181
Lutetium	Lu	71	9.85	17.76	8.12	84	1652
Magnesium	Mg	12	1.74	14.0	25.7	44	650
Manganese	Mn	25	7.43	7.39	22.6	198	1245
Mendelevium	Md	101	—	—	—	—	—
Mercury	Hg	80	13.55	14.81	61	—	-38
Molybdenum	Mo	42	10.22	9.39	4.98	328	2610
Neodymium	Nd	60	7.00	20.61	9.98	38	1019
Neon	Ne	10	—	—	—	—	-249
Neptunium	Np	93	—	—	27.5	100	637
Nickel	Ni	28	8.90	6.59	12.7	193	1453
Niobium	Nb	41	8.57	10.8	7.07	105	2468
Nitrogen	N	7	—	—	—	—	-210
Nobelium	No	102	—	—	—	—	—
Osmium	Os	76	22.57	8.43	4.7	540	2700
Oxygen	O	8	—	—	—	—	-219
Palladium	Pd	46	12.02	8.88	11.5	124	1552
Phosphorus	P	15	1.83	16.92	124	4.6	44

APPENDIX I (cont.)

Element	Symbol	Atomic number	Density, 10^3 kg m^{-3}	Molar volume, 10^{-8} m^3	Thermal expansion, 10^{-6} K^{-1}	Young's modulus, GN m^{-2}	Melting point, $^{\circ}\text{C}$
Platinum	Pt	78	21.45	9.09	8.95	170	1769
Plutonium	Pu	94	19.5	12.27	55	96.5	640
Polonium	Po	84	—	—	23	25.5	254
Potassium	K	19	0.86	45.47	83	3.5	64
Praseodymium	Pr	59	6.77	20.82	6.79	33	919
Promethium	Pm	61	—	—	9.0	42	1027
Protactinium	Pa	91	15.4	15.00	7.3	100	1230
Radium	Ra	88	5.0	45	20.2	16	700
Radon	Rn	86	—	—	—	—	−71
Rhenium	Re	75	21.04	8.85	6.63	460	3180
Rhodium	Rh	45	12.44	8.27	8.40	372	1966
Rubidium	Rb	37	1.53	55.87	88.1	2.7	39
Ruthenium	Ru	44	12.2	8.29	9.36	410	2500
Samarium	Sm	62	7.49	20.07	10.4	34	1072
Scandium	Sc	21	2.99	14.89	10.0	79	1539
Selenium	Se	34	4.79	16.48	36.9	58	217
Silicon	Si	14	2.33	12.06	3.07	103	1410
Silver	Ag	47	10.49	10.28	19.2	80.5	961
Sodium	Na	11	0.97	23.67	70.6	8.9	98
Strontium	Sr	38	2.60	34	20	13.5	768
Sulphur	S	16	2.07	15.5	64	19.5	119
Tantalum	Ta	73	16.6	10.9	6.55	181	2996
Technetium	Tc	43	—	—	8.06	370	2130
Tellurium	Te	52	6.24	20.45	16.77	41	450
Terbium	Tb	65	8.25	19.26	10.3	57.5	1356
Thallium	Tl	81	11.85	17.25	29.4	8	303
Thorium	Th	90	11.66	19.90	11.2	74	1750
Thulium	Tm	69	9.31	18.15	13.3	75	1545
Tin (gray)	Sn	50	7.30	16.26	5.3	52	232
Titanium	Ti	22	4.51	10.63	8.35	106	1668
Tungsten	W	74	19.3	9.53	4.59	396	3410
Uranium	U	92	19.07	12.48	12.6	186	1132
Vanadium	V	23	6.1	8.35	8.3	132	1900
Xenon	Xe	54	—	—	—	—	−112
Ytterbium	Yb	70	6.96	24.86	24.96	18	824
Yttrium	Y	39	4.47	19.89	12.0	65	1509
Zinc	Zn	30	7.13	9.17	29.7	92	420
Zirconium	Zr	40	6.49	14.06	5.78	92	1852

APPENDIX

II Properties of Engineering Materials

Material	Density, 10^3 kg m^{-3}	Thermal conductivity, $J, \text{ m}^{-1} \text{ K}^{-1} \text{ s}^{-1}$	Thermal expansion, 10^{-6} K^{-1}	Young's modulus, GN m^{-2}	Tensile strength, MN m^{-2}	% elongation
0.2% C Steel	7.86	50	11.7	210	350	30
0.4% C Steel	7.85	48	11.3	210	600	20
0.8% C Steel	7.84	46	10.8	210	800	8
18/8 stainless steel	7.93	15	9.0	210	700	65
Gray cast iron	7.15	—	10.4	210	250	—
White cast iron	7.7	—	9.0	210	—	—
Brass (70 Cu 30 Zn)	8.5	120	19.8	110	400	50
Bronze (95 Cu 5 Sn)	8.8	80	18.0	110	600	30
Al_2O_3	3.8	29	9.0	345	—	—
MgO	3.6	40	9.0	205	—	—
Silica glass	2.2	1.2	0.54	70	50	—
Vycor glass	2.2	1.2	0.63	70	50	—
Borosilicate glass	2.4	1.0	2.7	65	50	—
Graphite	1.9	40	5.4	8	—	—
Silicon carbide	3.17	12	4.5	350	—	—
Vulcanized rubber	1.2	0.12	80	1–3	—	—
Polyethylene	0.9	0.34	180	0.2	15–35	100–500
Polystyrene	1.05	0.08	60	3.0	55	3
Polytetra- fluoroethylene	2.2	0.20	100	0.5	15–25	200–300
Polymethyl methacrylate	1.2	0.20	90	3.5	55	5
Nylon	1.15	0.25	100	3.0	80	40–80

Index

- Acceptor level, 366
Activated complex, 17
Activation energy, 16
 for creep, 288
 for diffusion, 184, 193–194
 for dislocation motion, 271–272
Activation volume, 272–273
Age hardening, 222–223
Ageing, 221–223
Ag-Pt phase diagram, 155
Aircraft structural material, 222
Alclad, 185, 326
Al-Cu phase diagram, 220–221
Alkali metals, 58, 94
Alloy carbides, 219
Alloying elements in steel, 219
Alloy steels, 215–216, 219–220
Alnico alloys, 406
 $\text{Al}_2\text{O}_3\text{-Cr}_2\text{O}_3$ phase diagram, 152–153
Alumel, 319
Alumina, 67, 102, 152–153, 163, 379
Aluminium, 71, 94, 185, 318, 319, 326,
 365–366, 377
Aluminium alloys, 185, 220–222, 263
Amorphous semiconductors, 369
Anelastic behaviour, 249–252
Annealing of glass, 228
Annealing of steel, 215
Annealing twins, 141
Antiferromagnetism, 397, 399
Arrhenius equation, 16–17, 18, 184, 289,
 373
Asbestos, 105, 106
ASTM grain size, 279–281
Athermal obstacles, 273, 274
Atomic diameter, 29, 64, 65
Atomic mass unit, 81
Ausforming, 220
Austempering, 220
Austenite, 165, 168, 215–217
Avalanche breakdown, 375
Avogadro's number, 15

Bainite, 215, 218, 220
Bakelite, 107, 333, 418
Barium ferrite, 406
Barium titanate, 421–423
Basal plane, 36, 264, 275
Base, 379
Basis, 29, 85, 91
BCC crystals, 29, 94, 95, 97, 276
BCC space lattice, 25–26, 31
BCS theory, 345
B-H curve, 400–401
Binary phase diagrams, 151–156
Bohr magneton, 396
Boiling points, 67, 71, 74
Boltzmann's constant, 13, 15
Bond energy, 63–64, 69–71, 75–76
Bonding, 63–73
 and properties, 74–76
Bond length, 64
Boron, 189, 244, 365, 376
Bragg angle, 39, 358
Bragg law, 38, 356
Branching, 112, 244
Brasses, 164–165, 430
Bravais lattices, 25–29
Brillouin zones, 359–360
Brittle fracture, 300–303, 307–308
Bulk modulus, 242
Burgers circuits, 128–130
Burgers vectors, 128–133, 135, 270–272

Cadmium sulphide, 384
Capacitance, 413–414
Capacitor, 413, 423
Carbon nanotubes, 86–87

- Carburization, 186–187, 311
- Carnot cycle, 289
- Cast irons, 168–169, 251, 430
- Cathodic protection, 327
- Caustic embrittlement, 328
- C-curves, 213, 215–216, 219
- Cellulose, 110
- Cement, 106
- Cementite, 166–169, 214–215, 219
- Ceramics and glasses, 3
- Characteristic radiation, 37
- Charge coupled device, 383
- Chemical bonding (*see* Bonding)
- Chemical vapour deposition, 376–377
- Chromel, 319
- Chromium, 29, 92, 219–220, 318–319, 327–328
- Clay, 106
- Cleavage direction, 106
- Cleavage planes, 372, 379
- Climb of dislocations, 135–136, 288
- Close packing geometry, 88–91
- Coalescence, 222, 290
- Coarsening, 222, 290
- Coatings, 326–327
- Coercive field, 401, 406
- Coiled molecules, 247, 255
- Cold work, 228–229, 277–278, 285
- Collector, 383
- Collision time, 339
- Columnar crystals, 226
- Components, 149
- Composite materials, 244–245
- Concentration cell, 323–324
- Concrete, 106
- Conduction band, 360–366, 419
- Conductivity
 - electrical, 74, 165, 194, 231, 339–340, 363, 366–368
 - thermal, 74, 430
- Conductors, 333, 342, 360
- Configurational entropy, 13–15, 124, 247–248
- Constantan, 343
- Contact angle, 210
- Contact potential, 380
- Coordination number, 85, 91, 97
- Copolymers, 112,
- Copper, 71, 74, 76, 164–165, 220–221, 223, 269, 271, 273, 281–282, 342
- Coring, 158
- Corrosion, 185–186, 319–328
- Cotton, 110
- Cottrell atmospheres, 133, 282
- Coulomb attraction, 66
- Covalent bond, 64, 68–70, 75
- Covalent crystals, 83–88
- Cracks, 219, 300–309, 311
- Creep, 253, 287–290, 307–308
- Creep-resistant materials, 289–290
- Critical nucleus, 206
- Critical resolved shear stress, 265, 267
- Cross links, 111, 196, 247–248
- Cross slip, 136–137, 278, 287
- Crystal directions, 31–36
- Crystallinity, 111–112
- Crystallization, 224–226
- Crystal planes, 32–36
- Crystals
 - geometry of, 24–36
 - orientation, 372
 - structure determination of, 42–45
 - structure of, 7, 29–31, 83–106
- Crystal systems, 25–28
- Cubic crystals, 29–31, 85, 90, 100–101
- Cubic space lattices, 25–26
- Curie temperature, 298, 421–423
- Czochralski method, 370–372
- Damping capacity, 251–252
- Dashpot model, 252–255
- de Broglie wavelength, 334–335
- Debye Scherrer camera, 41
- Decarburization, 187–188, 311
- Deformation, 261–286
- Deformation twins, 141
- Degradation, 107
- Degree of polymerization, 108
- Degrees of freedom, 149–150
- Dendritic structure, 158–159
- Density, 75–76, 83, 86, 112, 120, 427–430
- Devices, 3
- Dezincification, 328
- Diamagnetism, 394–395
- Diamond, 70, 85–86, 271, 274, 360–361
- Diamond cubic structure, 44–45, 85–86
- Dielectric constant, 365, 413, 417–418
- Dielectric strength, 418–419
- Diffractometer, 42, 43
- Diffusion, 178–196
 - activation energy, 184, 193–195
 - coefficient, 179, 182–185, 192–193
 - interstitial, 190–192, 250–251
 - mechanisms of, 190–193
 - substitutional, 192–193

- Diffusional creep, 288
- Diodes, 379–382
- Dipole moment, 72, 415
- Dislocations, 126–137, 267–286, 371
 - climb of, 135–136, 288
 - cross slip of, 136–137, 278, 287
 - density of, 133, 277–278, 285
 - edge, 126–128
 - elastic energy of, 131–132
 - in cubic crystals, 132
 - movement of, 134–137, 269–286
 - multiplication of, 274–275
 - pile up of, 277, 278, 279
 - screw, 129–130
 - sources, 274–275
 - width of, 270
- Dispersion hardening, 290
- Domain structure, 399–401
- Donor level, 364–365
- Doping, 189, 364, 375
- Double bond, 69, 70
- d*-states, 54, 59–60, 75, 397–399
- Ductile-brittle transition, 305–307, 404
- Ductile fracture, 298–300, 307–308
- Ductility, 167
- Duralumin, 185–186, 220–222, 326
- Dynamic recovery, 278

- Ebonite, 111, 244
- Eddy current loss, 403–405
- Edge dislocation, 126–128
- Einstein relationship, 55
- Elastic anisotropy, 242
- Elastic behaviour, 239–252
- Elasticity, 239–252, 255
- Elastic moduli, 239–246, 427–430
- Elastomers, 110–111, 246–249
- Electrical conductivity, 74, 165, 194, 231, 339–340, 363, 366–368
- Electric breakdown, 419–420
- Electrode potential, 320–321
- Electromigration, 377
- Electron, 54
- Electron affinity, 61–62
- Electron compounds, 97
- Electron diffraction, 5
- Electronegativity, 62, 415
- Electronic polarization, 414, 416–417
- Electronic structure, 7, 334–338, 356–360
- Electron micrographs, 5, 279, 284, 299
- Electron microscope, 5–6
- Electron probability density, 56–58, 356–357
- Electron probe microanalyzer, 6
- Elements, 427–429
- Elinvars, 246
- Elongation, 164, 261–262, 277–278, 285, 430
- Embryos, 206
- Emitter, 383
- Endurance limit, 311
- Energy bands, 359–367
- Energy gap, 359–367
 - direct, 369
 - in compounds, 368
- Energy product, 406
- Engineering materials, 2–3, 430
- Engineering strain, 261
- Engineering stress, 261, 299
- Enthalpy, 12–13
 - of atomization, 64, 67
 - of formation, 124, 125
 - of fusion, 12
 - of motion, 191, 192
- Entropy, 12–15
 - configurational, 14–15, 124, 247–248
 - thermal, 12, 15
- Epitaxial growth, 376–377
- Epoxy, 245
- Equilibrium configurations, 11
- Equilibrium diagrams (*see* Phase diagrams)
- Error function, 181–182
- ESD magnets, 406
- Etch pits, 133
- Ethylene, 107–109
- Eutectic mixture, 157–158, 160–161
- Eutectic phase diagram, 153–154
- Eutectic reaction, 154, 156, 165
- Eutectic temperature, 154–155
- Eutectoid reaction, 155–156
- Eutectoid steel, 166, 214–215
- Exchange energy, 397–398
- Excimer lasers, 374
- Exhaustion region, 367–368
- Explosive forming, 348
- Extinction rules, 44–45
- Extrinsic semiconductors, 364–368

- Face centred cubic stacking, 89–90
- Fatigue fracture, 310–311
- Fatigue resistance, 187, 310–311
- FCC crystals, 29–30, 90, 276
- FCC space lattice, 25–26

- FCC stacking, 89–90
- Fe, 64, 71, 94, 121–122, 268, 271, 273–274, 402–404
- Fe-Fe₃C phase diagram, 165
- FeO, 123, 194
- Fe₃O₄, 399
- Fe phase diagram, 150–151
- Fermi–Dirac statistics, 337, 362
- Fermi energy (level), 337–338, 362, 379–380
- Ferrimagnetism, 398–399
- Ferrite phase, 166–167, 214–215, 282
- Ferrites, 102, 399, 405, 406
- Ferroelectric domains, 422
- Ferroelectric materials, 420–423
- Ferromagnetism, 395, 397–398
- Fibre reinforcement, 244–245, 252
- Fibres, 2, 109–110
- Fick's first law, 179–180
- Fick's second law, 180–182
 - applications based on, 182–189
 - solution to, 181–182
- Field-ion micrograph, 6, 122
- Field-ion microscope, 6
- Float zone method, 370, 372
- Forbidden gap, 359–361
- Forward bias, 380–382
- Fourier's law, 179
- Four parameter model, 255
- Fracture, 261–262, 298–311
 - brittle intergranular, 307–308
 - mechanism maps, 307–308
- Fracture stress, 302
- Fracture toughness, 304–305
- Frank-Read source, 274–275, 276, 277
- Free electrons, 71, 74
 - conduction by, 338–340
- Free electron theory, 334–338
- Free energy, 12–13
- Frenkel defect, 122–123, 125, 193
- Fullerenes, 86

- Gallium arsenide, 368, 369, 383–384
- Galvanic cell, 322–324
- Galvanic protection, 326
- Galvanic series, 321
- Galvanized iron, 326
- Garnets, 402, 405
- Germanium, 70, 87, 361, 368, 369
- Gibbs free energy, 13, 124
- Gibbs phase rule, 148, 161
- Glass transition, 226–228
 - temperature, 227–228

- Glide of dislocation, 135
- G-P zones, 223
- Grain boundary, 138, 278–280
 - sliding, 288–289
- Grain growth, 141, 231
- Grain size, 225–226, 278–280, 286
 - ASTM number, 279–281, 286
- Graphite, 84–85, 86
 - flakes, 169, 251
- Gray cast iron, 169, 251, 430
- Griffith criterion, 300–304
- Ground state, 55
- Growth kinetics, 211–212
- Growth of crystals, 158, 204, 224, 370–372

- Hall-Petch equation, 279, 286, 306
- Halogens, 83
- Hardenability, 219
- Hard magnets, 405–406
- Hardness of steel phases, 218
- HCP crystal, 90–91, 92, 94, 275–276
- HCP stacking, 90–92, 140
- Heating elements, 88, 343
- Heat of fusion, 12–13, 224
- Heat treatment, 214–223, 225, 309
- Heisenberg principle, 54
- Heterogeneous nucleation, 209–211, 224
- Heusler alloys, 398
- Hexagonal close packing, 90–92
- Hexagonal unit cell, 27, 29, 36
- High angle boundaries, 139
- High speed steel, 219
- High temperature materials, 289–290
- Holes, 363, 366
- Homogeneous nucleation, 205–209
- Hooke's law, 239, 252
- Hume-Rothery's rules, 96, 152, 282, 286
- Hund's rule, 58, 397
- Hybridized orbitals, 68
- Hydrogen atom, 55–57
- Hydrogen bond, 64, 72, 110, 244
- Hydrogen electrode, 320
- Hydrogen molecule, 68
- Hysteresis loop, 251, 400–402

- Ice, 73, 83
- Icosahedron, 86–87
- Ideal (perfect) crystals, 120, 266–268
- Immobile dislocation, 277
- Impact test, 305–306
- Imperfections in crystals, 120–142

- Inconel, 319
- Induction, 394
- Inert gas configuration, 58, 66
- Inert gases, 73
- Inhibitors, 326
- Insulators, 333, 361
- Integrated circuits, 369–379
- Interatomic forces, 63, 239–241
- Interface, 139
- Intergranular failure, 307–308, 309, 327–328
- Intermediate structures, 97
- Intermetallic compounds, 97
- Internal energy, 12
- Interphase boundary, 139
- Interplanar spacing, 36, 39
- Interstitial diffusion, 190–192
- Interstitial impurity, 121–122
- Interstitial voids, 93, 121–122
- Interstitialcy mechanism, 191
- Intrinsic semiconductors, 361–364
- Invar, 246
- Invariant reaction, 155–156
- Invariant temperature, 155, 162
- Ion exchange, 309
- Ion implantation, 376
- Ionic bond, 64, 65–67
- Ionic conductivity, 194, 334
- Ionic crystals, 97–102
- Ionic diameter, 64
- Ionization potential, 60–61
- Iron (*see* Fe)
- Isoprene, 110

- Josephson junction, 347
- Junction, 379–384

- K_{α} radiation, 38
- Kanthal, 319, 342, 343
- Kinetics, 16–18
- Kirchhoff's law, 131
- Kirkendall effect, 189–190

- Laser, 102, 383–384
- Lattice friction, 273
- Lattice parameter, 29
- Lattice vibration frequency (*see* Vibrational frequency)
- Laue technique, 40
- Lead, 71, 94, 154, 171, 287

- Levels of structure, 3–7
- Lever rule, 159–162
- Ligancy, 97–99
- Line imperfections, 120, 126–137
- Liquidus, 153, 154
- Lithium, 222–223
- Long-chain polymers (*see* Polymers)
- Loss angle, 418
- Low angle boundaries, 139–140

- Machines, 3
- Macromolecule, 107
- Macrostructure, 4
- Madelung constant, 67
- Magnesia, 67, 163
- Magnetite, 399
- Magnetization, 394, 396, 400
- Majority carriers, 365, 366
- Malleable cast iron, 169
- Manganin, 342, 343
- Maraging steel, 285
- Martempering, 220
- Martensite, 215–220, 282, 323
- Mask, 374, 378–379
- Materials science and engineering, 1–2
- Mattheissen rule, 340–341
- Maxwell-Boltzmann statistics, 16, 206
- Maxwell element, 252–254
- Mean free path, 340
- Mechanical properties, 75, 238–290, 430
- Meissner effect, 344–345, 346
- Melting points, 67, 70, 71, 427–430
- Metallic bond, 71, 75, 271
- Metallic crystals, 88–97, 271, 273–274, 360
- Metallic glass, 203, 404
- Metallization, 377
- Metals and alloys, 2, 88–97
- Metastability, 10–11
- Metastable, 10, 11, 165, 219, 282
- MHD generators, 347
- Mica, 105, 106, 333, 418, 419
- Microalloyed steels, 280–281
- Microconstituent, 157, 215
- Microstructure, 4–5, 141, 157, 158, 166, 168, 169
- Mild steel, 133, 168, 282–283, 430
- Miller-Bravais indices, 36
- Miller indices, 31–36
- Miniaturization, 379
- Mobility, 363, 368
- Modulus, elastic, 240–243

- Molecular beam epitaxy, 376–377
 Molecular crystal, 30
 Monoatomic crystals, 29–30
 Monomer, 106, 109, 110
 MOSFET, 383
- NaCl structure, 100–101, 133
 Natural rubber, 110
 $\text{Nd}_2\text{Fe}_{14}\text{B}$, 406
 Necking, 261–262, 287, 298–299
 Nernst–Einstein relation, 194
 Neutrons, 54
 Newton's law of viscous flow, 252
 Nichrome, 319, 342, 243
 Nickel, 71, 94, 215, 219, 268, 273, 282, 290
 Nitriding, 186
 Nodular iron, 169
 Noncrystalline state, 82–83, 226
 Nonsteady state flow, 180
 Normalized steel, 215
n-p-n junction, 382–383
n-type semiconductors, 365
 Nuclear structure, 7
 Nucleation and growth, 204
 Nucleation kinetics, 205–211
 Nucleus, 54
 in nucleation, 206
 Nylon, 109–110, 228, 418, 430
- Octahedral coordination, 99
 Octahedral voids, 93–94
 Octahedron, 94
 Ohm's law, 179, 339
 Opaque glasses, 223
 Orbitals, 54–60
 Ordered state, 97
 Order of reflection, 39
 Orlon, 109
 Orthorhombic cell, 26, 28
 Overageing, 222, 283, 286
 Oxidation, 316–319, 373
 Oxygen molecule, 69
- Packing efficiency, 86, 91–92
 Paramagnetism, 395
 Partial dislocation, 132
 Passivation, 325, 327
 Pauli exclusion principle, 55, 66, 73, 335, 347
- Pauling's electronegatives, 62
 Pb–Sn phase diagram, 154
 Pearlite, 166–167, 214–215, 218
 Peierls–Nabarro (P–N) stress, 270–271, 272, 273
 Perfect crystal, 266–267
 Periodic table, 55–60
 Peritectic reaction, 156, 164, 165
 Peritectic system, 155–156
 Peritectoid reaction, 156
 Permalloy, 405
 Permanent magnets, 405–406
 Permeability, 394
 Permittivity, 413
 Perovskite, 347
 Petch equation, 279
 Phase, 149
 Phase rule, 148–149
 Phase transformations, 201–231
 applications, 213–231
 free energy change in, 202–203
 overall kinetics, 211–213
 time scale, 203
 Phenolformaldehyde, 244
 Phonons, 340
 Phosphates, 225, 328
 Photoconductor, 384
 Photolithography, 374–375
 Photon detectors, 384
 Photoresists, 374
 Pi-bonds, 69
 Piezoelectric device, 104, 251, 423
 Pilling–Bedworth ratio, 316
 Pinning action, 230
 Pipe diffusion, 194
 Pitting, 326
 Planck's constant, 55, 334
 Plastic deformation, 260–286
 Plasticizers, 112, 228
 Plastics, 2, 107
 Platinum, 6, 71, 343
p-n junction, 379–382
 Point imperfections, 120–125
 Poisson's ratio, 242
 Polarization, 324–325, 415–417, 422
 Polyacrylonitrile, 109, 228
 Polycrystalline, 23, 138
 Polyester fibres, 110, 228
 Polyethylene, 108–109, 228, 244, 333, 418, 430
 Polyisoprene, 110–111, 228
 Polymers, 2, 106–111, 228, 244, 245, 251, 307

- Polymethylmethacrylate, 109, 430
- Polypropylene, 109, 228
- Polystyrene, 109, 228, 430
- Polytetrafluoroethylene, 109, 430
- Polyvinylchloride, 109, 228, 418
- Porcelain, 333, 418
- Portland cement, 106
- Potential energy, 10–11, 63, 239, 240, 266–267, 356
- Powder method, 41–42
- Precipitates, 221–223
- Precipitation, 220–223, 283–284
- Precipitation hardening, 283–284, 286
- Primary bonds, 64, 74, 244
- Prismatic plane, 36, 264, 276
- Properties
 - of elements, 427–429
 - of engineering materials, 430
 - structure correlation, 7
- Protons, 54
- p*-states, 54–55, 57–60, 396
- p*-type semiconductors, 366
- Pyramidal planes, 264, 276
- Pyrex, 104
- Pyroceram, 225–226, 309

- Quantum numbers, 54–55, 336
- Quantum states, 54–55, 336
- Quartz, 104, 105, 251

- Radius ratio, 98–101
- Rapid solidification processing, 224
- Rare earth elements, 60, 398, 406
- Rate of a reaction, 16–18
- Rayon, 110
- Recovery, 229, 231, 278
- Recrystallization, 138, 141, 229–230
 - temperature, 230
- Rectifier, 382
- Refractive index, 223
- Refractory, 67, 170, 289, 308
- Relaxation processes, 249–252
- Relaxation time, 249–252
- Residual induction, 400, 405–406
- Residual stresses, 219, 309
- Resistivity range, 333–334
- Resistors, 343
- Resolution, 3–5
- Resolved shear stress, 265
- Resonance, 418
- Resonating bond, 84
- Reverse bias, 381–383

- Ring mechanism, 191
- Rockwell hardness, 218, 219
- Rotating crystal method, 40
- Rubber, 110–111, 333, 418, 430
 - elasticity, 246–249
- Ruby, 102

- Scanning electron microscope, 5
- Schottky defect, 122, 125, 193
- Screw dislocations, 129–131, 278, 287
 - cross slip of, 136–137, 278, 287
- Season cracking, 328
- Secondary bonds, 64, 72–73, 74, 244
- Seeding, 210
- Semiconductor(s), 87, 333, 355–369, 379–384
 - devices, 379–384
 - doping of, 189, 364–365, 375–376
 - extrinsic, 364–368
 - intrinsic, 361–364
 - materials, 368–369
- Sessile dislocation, 277
- SG iron, 169
- Shear modulus, 131, 242, 267
- Shear strain, 252–254, 267
- Shear strength, 267
- Shear stress, 265–268
- Shot peening, 311
- Sigma bond, 69
- Silica, 103–104, 105
- Silicate
 - glasses, 225, 226–227
 - structural units, 105
 - structures, 104–106
 - tetrahedron, 103
- Silicon, 70, 87, 271, 273, 308, 363–366, 369–370
 - doping of, 189, 364, 365, 375
- Silicon carbide, 70, 87–88, 430
- Silicon nitride, 289
- Simple cubic crystal, 30, 136
- Simple cubic space lattice, 25–26
- Single crystals, 23, 138, 224, 370–372
 - creep of, 290
 - growth of, 224, 370–372
 - semiconductor, 370–372
- Slip, 134–135, 263–265
- Slip directions, 264
- Slip lines, 264
- Slip plane, 127, 264
- Slip systems, 264
- S-N curves, 312

- Soda-lime glass, 104, 226, 334
- Sodium chloride, 100–101, 132–133, 244, 264, 271
- Soft magnetic materials, 402–405
- Solder alloy, 171, 325
- Solidification, 157–158, 224
- Solid solutions, 96–97, 102, 281–282, 286
- Solid state devices, 87, 224, 355, 379–384
- Solidus, 153, 154
- Solute atoms, 121–122, 281–282, 286, 340
- Solute drag effect, 230, 231
- Solvus, 154
- Space lattices, 24–31
 - definition, 24
- sp^2 bonds, 84
- sp^3 bonds, 68, 85, 100
- Spheroidal graphite iron, 169
- Spinel, 102, 405
- Spin of electrons, 55, 395–396
- Splat cooling, 203, 224
- Spring-dashpot models, 252–255
- Sputtering, 375
- Stability, 10–11
- Stable equilibrium, 11
- Stacking faults, 140
- Stainless steels, 263, 273, 318–319, 327–328, 430
- Standard potential, 320
- Steady state flow, 179–180
- Steels,
 - phase diagram, 165–167
 - properties, 430
 - transformations in, 214–220
 - uses of, 168
- Stiffness, 243
- Strain energy, 131, 208, 216
- Strain hardening (*see* Work hardening)
- Strain rate, 262, 252–254, 271–272
- Strain rate sensitivity, 262
- Strength coefficient, 262
- Strength of materials, data, 218, 285, 430
- Stress cells, 324
- Stress concentration, 279, 303, 311
- Stress corrosion, 328
- Stress intensity factor, 305
- Stress-strain curve, 249, 250, 261–262, 277
- Structure
 - determination of, 42–45
 - levels of, 3–7
 - property relationships, 7–8
- Structures, 3
- Structure-sensitive property, 120
- Styrofoam, 109
- Substitutional impurity, 121
- Substructure, 5–6
- Superalloy, 290
- Supercomputers, 383
- Superconductors, 346–348
- Supercooling, 203–206
- Supermalloy, 405
- Superplastic 262–263
- Surface hardening, 186–187
- Surface
 - energies, 137–138
 - imperfections, 137–141
- Surfaces, 137–138
- Susceptibility, 394–395
- Symmetry, 28
- Talc, 106
- Tannins, 328
- Teflon, 109
- Tempering, 219, 309, 323
- Tensile strength, 261, 430
- Tensile stress-strain curve, 261–263
- Terylene, 110
- Tetragonal martensite, 217–218
- Tetragonal unit cell, 26, 28
- Tetrahedral angle, 68, 85
- Tetrahedral coordination, 85, 103
- Tetrahedral void, 93
- Tetrahedron, 68
- Texture, 402
- Thermal conductivity, 430
- Thermal energy, 11, 18, 179, 208, 212, 272–274, 287, 316, 362, 398, 419
- Thermal expansion, 65, 75–76, 246, 427–430
- Thermal stresses, 203, 219
- Thermodynamic functions, 12–14
- Thermoelectric potential, 343
- Thermometers, 228, 343
- Thermoplasts, 107
- Thermosets, 107
- Thoria, 290, 343
- Tie line, 152–153, 159, 161–162
- Tilt boundary, 139–140, 229
- Titanium, 76, 94, 290
- TiB₂, 244
- Tool steels, 219
- Transistor, 379, 382–383
- Transition metals, 59, 75–76, 271, 274, 397–398

- Transition temperature
 - ductile-brittle, 305, 404
 - glass, 227
 - superconducting, 344
- Translucent glass, 223
- Tremolite, 105
- Trichlorosilane, 370
- Triple bond, 70
- Triple points, 151
- True strain, 261
- True stress, 261, 299
- T-T-T diagrams, 213–214, 216
- Tungsten, 71, 289, 343
- Tungsten carbide, 219
- Turbine blades, 289, 290
- t vector, 128–131
- Twin boundary, 140–141, 278–279
- Twinning, 263–264
- Twins, 141
- Twist boundary, 140, 229

- Ultimate tensile strength, 261
- Unit cells, 7, 24–29, 85, 89–91, 100–101
- Unstable equilibrium, 11, 267

- Vacancy, 121
 - enthalpy of formation of, 124
- Vacancy diffusion, 190
- Valence band, 360–366
- van der Waals bond, 64, 73, 244
- Vibrational frequency, 15, 18, 191–193, 207
- Viscoelastic behaviour, 252–255
- Viscosity, 226–227, 252, 289

- Viscous flow, 252–255, 289
- VLSI, 369, 384
- Voids in close packing, 93–94
- Voigt-Kelvin element, 254–255
- Von Mises criterion, 264
- Vortex region, 346
- Vulcanization, 111

- Wafer, 372–373, 376, 378–379
- Water molecule, 68, 72, 83
- Wave form, 335–337, 356–357
- Wave number, 334–335, 356–360
- Whiskers, 269
- White cast iron, 168, 430
- White radiation, 37
- Wood, 110
- Work hardening, 261, 275–277, 285

- X-rays
 - diffraction by, 37–41, 263
 - wavelengths of, 37–38

- $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, 347–348
- Yield point, 261, 282–283
- Yield strength, 261, 285, 306
- Young's modulus, 240–243, 300–302, 427–430
- Yttrium-iron-garnet, 405

- Zinc, 71, 94, 164, 268, 281–282
- ZnO, 87, 123, 194
- ZnS, 87, 100
- Zone refining, 170, 372

ATOMIC AND IONIC RADII

(Units are in Å)

$$1 \text{ \AA} = 10^{-10} \text{ m} = 0.1 \text{ nm}$$

[illegible]

CRYSTAL STRUCTURES OF THE ELEMENTS

1. The structures are at room temperature, except in the case of the eighth column elements.
2. The a lattice parameter is shown for all cubic crystals.
3. For noncubic crystals, the c parameter is shown below the a parameter.

(Units are in Å)
1 Å = 10⁻¹⁰ m = 0.1 nm

CRYSTAL STRUCTURES OF THE ELEMENTS																	He HCP 3.57 5.83	
H	1. The structures are at room temperature, except in the case of the eighth column elements. 2. The a lattice parameter is shown for all cubic crystals. 3. For noncubic crystals, the c parameter is shown below the a parameter. (Units are in Å) $1\text{ Å} = 10^{-10}\text{ m} = 0.1\text{ nm}$																	Ne FCC 4.46
Li BCC 3.49	Be HCP 2.27 3.59											B Rhom.	C DC 3.57 (Hex.)	N	O	F	Ar FCC 5.31	
Na BCC 4.22	Mg HCP 3.21 5.21											Al FCC 4.05	Si DC 5.43	P Com- plex	S Com- plex	Cl		
K BCC 5.22	Ca FCC 5.58	Sc HCP 3.31 5.27	Ti HCP 2.95 4.68	V BCC 3.03	Cr BCC 2.88	Mn Cubic	Fe BCC 2.87	Co HCP 2.51 4.07	Ni FCC 3.52	Cu FCC 3.61	Zn HCP 2.66 4.95	Ga Com- plex	Ge DC 5.66	As Rhom.	Se Hex.	Br	Kr FCC 5.64	
Rb BCC 5.58	Sr FCC 6.08	Y HCP 3.65 5.73	Zr HCP 3.23 5.15	Nb BCC 3.30	Mo BCC 3.15	Tc HCP 2.74 4.40	Ru HCP 2.71 4.28	Rh FCC 3.80	Pd FCC 3.89	Ag FCC 4.09	Cd HCP 2.98 5.62	In Tetr. 3.25 4.95	Sn BCT (DC)	Sb Rhom.	Te Hex.	I Com- plex	Xe FCC 6.13	
Cs BCC 6.04	Ba BCC 5.02	La Hex.	Hf HCP 3.19 5.05	Ta BCC 3.30	W BCC 3.16	Re HCP 2.76 4.46	Os HCP 2.74 4.32	Ir FCC 3.84	Pt FCC 3.92	Au FCC 4.08	Hg	Tl HCP 3.46 5.52	Pb FCC 4.95	Bi Rhom.	Po SC 3.34	At	Rn	
Fr	Ra BCC 5.15	Ac FCC 5.31											B	C	N	O	F	Lu HCP 3.50 5.55
			Ce FCC 5.16	Pt Hex.	Nd Hex.	Pm	Sm Com- plex	Bu BCC 4.58	Gd HCP 3.63 5.78	Tb HCP 3.60 5.70	Dy HCP 3.59 5.65	Ho HCP 3.58 5.62	Er HCP 3.56 5.59	Tm HCP 3.54 5.56	Yb FCC 5.48	No	Lw	
			Th FCC 5.08	Pa Tetr. 3.92 3.22	U Com- plex	Np Com- plex	Pu Com- plex	Am Hex.	Cm Hex.	Bk Hex.	Cf Hex.	Es Hex.	Fm	Md				

Materials Science and Engineering



A First Course

FIFTH EDITION

V. Raghavan

This well-established and widely adopted book, now in its Fifth Edition, provides a thorough analysis of the subject in an easy-to-read style. It analyzes, systematically and logically, the basic concepts and their applications to enable the students to comprehend the subject with ease.

The book begins with a clear exposition of the background topics in chemical equilibrium, kinetics, atomic structure and chemical bonding. Then follows a detailed discussion on the structure of solids, crystal imperfections, phase diagrams, solid-state diffusion and phase transformations. This provides a good insight into the structural control necessary for optimizing the various properties of materials. The mechanical properties covered include elastic, anelastic and viscoelastic behaviour, plastic deformation, creep and fracture phenomena. The last four chapters are devoted to a detailed description of electrical conduction, superconductivity, semiconductors, and magnetic and dielectric properties.

KEY FEATURES

- All relevant units and constants listed at the beginning of each chapter
- A note on SI units and a full table of conversion factors at the beginning
- Sections on the new oxide superconductors and the fabrication of integrated circuits
- Examples with solutions and problems with answers
- Over 330 multiple choice questions with answers

This highly accessible and student-friendly text not only provides a masterly analysis of all the relevant topics, but also makes them comprehensible to the students through the skilful use of clear-cut diagrams, illustrative tables, worked-out examples, and in many other ways. It is a **must-have book** for all engineering and applied science students.

ABOUT THE AUTHOR

V. RAGHAVAN, Ph.D. (Sheffield), formerly Professor of Materials Science at the Indian Institute of Technology Delhi, taught materials science courses at Massachusetts Institute of Technology, USA, and at IIT Kanpur and IIT Delhi for many years. A Fellow of the American Society of Metals and an Honorary Member of the Indian Institute of Metals, Professor Raghavan is the recipient of the Brunton Medal of Sheffield University, UK (1964), the G.D. Birla Gold Medal of Indian Institute of Metals (1985), and the Hume-Rothery Award of the Institute of Materials London (2001).

