

GLOBAL  
EDITION



# Solid State Electronic Devices

SEVENTH EDITION

Ben G. Streetman • Sanjay Kumar Banerjee

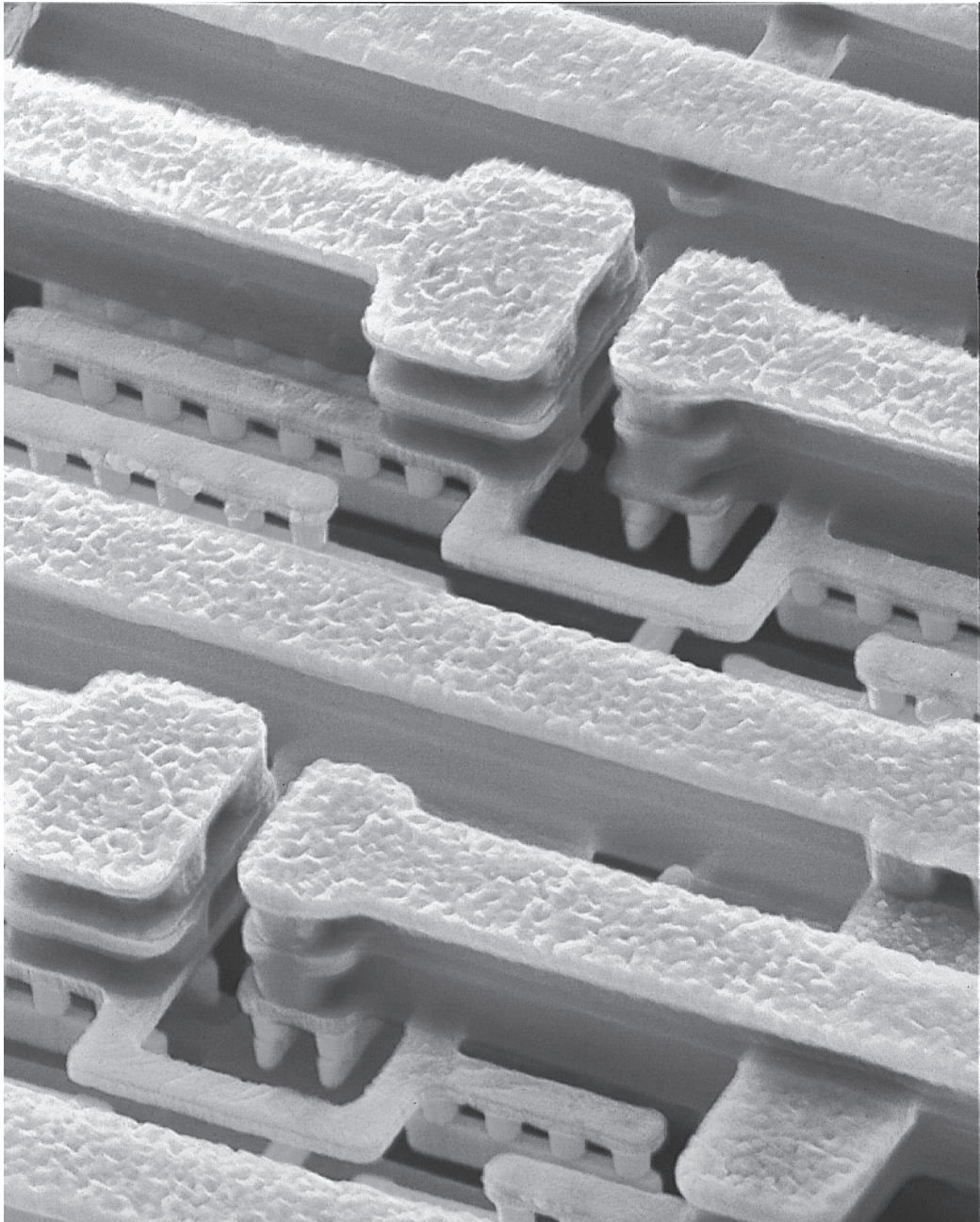
ALWAYS LEARNING

PEARSON

---

# Solid State Electronic Devices





**Multilevel copper metallization of a complementary metal oxide semiconductor (CMOS) chip.** This scanning electron micrograph (scale: 1 cm = 3.5 microns) of a CMOS integrated circuit shows six levels of copper metallization that are used to carry electrical signals on the chip. The inter-metal dielectric insulators have been chemically etched away here to reveal the copper interconnects. (Photograph courtesy of IBM.)

---

SEVENTH EDITION GLOBAL EDITION

# Solid State Electronic Devices

**BEN G. STREETMAN AND SANJAY KUMAR BANERJEE**

*Microelectronics Research Center  
Department of Electrical and Computer Engineering  
The University of Texas at Austin*

PEARSON

Boston Columbus Indianapolis New York San Francisco Hoboken  
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto  
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Vice President and Editorial Director, ECS: *Marcia J. Horton*  
Senior Editor: *Andrew Gilfillan*  
Head of Learning Asset Acquisitions, Global Editions: *Laura Dent*  
Assistant Acquisitions Editor, Global Editions: *Aditee Agarwal*  
Executive Marketing Manager: *Tim Galligan*  
Marketing Assistant: *Jon Bryant*  
Permissions Project Manager: *Karen Sanatar*

Senior Managing Editor: *Scott Disanno*  
Production Project Manager: *Greg Dulles*  
Project Editor, Global Editions: *Donald Villamero*  
Operations Specialist: *Linda Sager*  
Senior Production Manufacturing Controller, Global Editions: *Trudy Kimber*  
Cover Designer: *Lumina Datamatics Ltd.*  
Cover Image: *RaJi/Shutterstock*  
Media Production Manager, Global Editions: *Vikram Kumar*

Pearson Education Limited  
Edinburgh Gate  
Harlow  
Essex CM20 2JE  
England

and Associated Companies throughout the world

Visit us on the World Wide Web at:  
[www.pearsonglobaleditions.com](http://www.pearsonglobaleditions.com)

© Pearson Education Limited 2016

The rights of Ben G. Streetman and Sanjay Kumar Banerjee to be identified as the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

*Authorized adaptation from the United States edition, entitled Solid State Electronic Devices, ISBN 978-0-13-335603-8, by Ben G. Streetman and Sanjay Kumar Banerjee, published by Pearson Education © 2015.*

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a license permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

10 9 8 7 6 5 4 3 2 1

ISBN 10:1-292-06055-7

ISBN 13: 978-1-292-06055-2

Typeset by Jouve India in 10/12 Times Ten LT Std Roman.  
Printed and bound in Great Britain by Courier Kendallville

PREFACE 13

ABOUT THE AUTHORS 17

## 1

CRYSTAL PROPERTIES AND GROWTH  
OF SEMICONDUCTORS 21

- 1.1 Semiconductor Materials 21
- 1.2 Crystal Lattices 23
  - 1.2.1 Periodic Structures 23
  - 1.2.2 Cubic Lattices 25
  - 1.2.3 Planes and Directions 27
  - 1.2.4 The Diamond Lattice 29
- 1.3 Bulk Crystal Growth 32
  - 1.3.1 Starting Materials 32
  - 1.3.2 Growth of Single-Crystal Ingots 33
  - 1.3.3 Wafers 35
  - 1.3.4 Doping 36
- 1.4 Epitaxial Growth 37
  - 1.4.1 Lattice-Matching in Epitaxial Growth 38
  - 1.4.2 Vapor-Phase Epitaxy 40
  - 1.4.3 Molecular Beam Epitaxy 42
- 1.5 Wave Propagation in Discrete, Periodic Structures 44

## 2

ATOMS AND ELECTRONS 52

- 2.1 Introduction to Physical Models 53
- 2.2 Experimental Observations 54
  - 2.2.1 The Photoelectric Effect 54
  - 2.2.2 Atomic Spectra 56
- 2.3 The Bohr Model 57
- 2.4 Quantum Mechanics 61
  - 2.4.1 Probability and the Uncertainty Principle 61
  - 2.4.2 The Schrödinger Wave Equation 63
  - 2.4.3 Potential Well Problem 65
  - 2.4.4 Tunneling 68
- 2.5 Atomic Structure and the Periodic Table 69
  - 2.5.1 The Hydrogen Atom 70
  - 2.5.2 The Periodic Table 72

## 3

**ENERGY BANDS AND CHARGE CARRIERS IN SEMICONDUCTORS 83**

- 3.1 Bonding Forces and Energy Bands in Solids 83
  - 3.1.1 Bonding Forces in Solids 84
  - 3.1.2 Energy Bands 86
  - 3.1.3 Metals, Semiconductors, and Insulators 89
  - 3.1.4 Direct and Indirect Semiconductors 90
  - 3.1.5 Variation of Energy Bands with Alloy Composition 92
- 3.2 Charge Carriers in Semiconductors 94
  - 3.2.1 Electrons and Holes 94
  - 3.2.2 Effective Mass 99
  - 3.2.3 Intrinsic Material 103
  - 3.2.4 Extrinsic Material 104
  - 3.2.5 Electrons and Holes in Quantum Wells 107
- 3.3 Carrier Concentrations 109
  - 3.3.1 The Fermi Level 109
  - 3.3.2 Electron and Hole Concentrations at Equilibrium 112
  - 3.3.3 Temperature Dependence of Carrier Concentrations 117
  - 3.3.4 Compensation and Space Charge Neutrality 119
- 3.4 Drift of Carriers in Electric and Magnetic Fields 120
  - 3.4.1 Conductivity and Mobility 120
  - 3.4.2 Drift and Resistance 125
  - 3.4.3 Effects of Temperature and Doping on Mobility 126
  - 3.4.4 High-Field Effects 129
  - 3.4.5 The Hall Effect 129
- 3.5 Invariance of the Fermi Level at Equilibrium 131

## 4

**EXCESS CARRIERS IN SEMICONDUCTORS 142**

- 4.1 Optical Absorption 142
- 4.2 Luminescence 145
  - 4.2.1 Photoluminescence 146
  - 4.2.2 Electroluminescence 148
- 4.3 Carrier Lifetime and Photoconductivity 148
  - 4.3.1 Direct Recombination of Electrons and Holes 149
  - 4.3.2 Indirect Recombination; Trapping 151
  - 4.3.3 Steady State Carrier Generation; Quasi-Fermi Levels 154
  - 4.3.4 Photoconductive Devices 156
- 4.4 Diffusion of Carriers 157
  - 4.4.1 Diffusion Processes 158
  - 4.4.2 Diffusion and Drift of Carriers; Built-in Fields 160
  - 4.4.3 Diffusion and Recombination; The Continuity Equation 163

- 4.4.4 Steady State Carrier Injection; Diffusion Length 165
- 4.4.5 The Haynes–Shockley Experiment 167
- 4.4.6 Gradients in the Quasi-Fermi Levels 170

## 5 JUNCTIONS 179

- 5.1 Fabrication of p-n Junctions 179
  - 5.1.1 Thermal Oxidation 180
  - 5.1.2 Diffusion 181
  - 5.1.3 Rapid Thermal Processing 183
  - 5.1.4 Ion Implantation 184
  - 5.1.5 Chemical Vapor Deposition (CVD) 187
  - 5.1.6 Photolithography 188
  - 5.1.7 Etching 191
  - 5.1.8 Metallization 193
- 5.2 Equilibrium Conditions 194
  - 5.2.1 The Contact Potential 195
  - 5.2.2 Equilibrium Fermi Levels 200
  - 5.2.3 Space Charge at a Junction 200
- 5.3 Forward- and Reverse-Biased Junctions; Steady State Conditions 205
  - 5.3.1 Qualitative Description of Current Flow at a Junction 205
  - 5.3.2 Carrier Injection 209
  - 5.3.3 Reverse Bias 218
- 5.4 Reverse-Bias Breakdown 220
  - 5.4.1 Zener Breakdown 221
  - 5.4.2 Avalanche Breakdown 222
  - 5.4.3 Rectifiers 225
  - 5.4.4 The Breakdown Diode 228
- 5.5 Transient and A-C Conditions 229
  - 5.5.1 Time Variation of Stored Charge 229
  - 5.5.2 Reverse Recovery Transient 232
  - 5.5.3 Switching Diodes 236
  - 5.5.4 Capacitance of p-n Junctions 236
  - 5.5.5 The Varactor Diode 241
- 5.6 Deviations from the Simple Theory 242
  - 5.6.1 Effects of Contact Potential on Carrier Injection 243
  - 5.6.2 Recombination and Generation in the Transition Region 245
  - 5.6.3 Ohmic Losses 247
  - 5.6.4 Graded Junctions 248
- 5.7 Metal–Semiconductor Junctions 251
  - 5.7.1 Schottky Barriers 251
  - 5.7.2 Rectifying Contacts 253



- 5.7.3 Ohmic Contacts 255
- 5.7.4 Typical Schottky Barriers 257
- 5.8 Heterojunctions 258

## 6

### FIELD-EFFECT TRANSISTORS 277

- 6.1 Transistor Operation 278
  - 6.1.1 The Load Line 278
  - 6.1.2 Amplification and Switching 279
- 6.2 The Junction FET 280
  - 6.2.1 Pinch-off and Saturation 281
  - 6.2.2 Gate Control 283
  - 6.2.3 Current–Voltage Characteristics 285
- 6.3 The Metal–Semiconductor FET 287
  - 6.3.1 The GaAs MESFET 287
  - 6.3.2 The High Electron Mobility Transistor (HEMT) 288
  - 6.3.3 Short Channel Effects 290
- 6.4 The Metal–Insulator–Semiconductor FET 291
  - 6.4.1 Basic Operation and Fabrication 291
  - 6.4.2 The Ideal MOS Capacitor 295
  - 6.4.3 Effects of Real Surfaces 306
  - 6.4.4 Threshold Voltage 309
  - 6.4.5 MOS Capacitance–Voltage Analysis 311
  - 6.4.6 Time-Dependent Capacitance Measurements 315
  - 6.4.7 Current–Voltage Characteristics of MOS Gate Oxides 316
- 6.5 The MOS Field-Effect Transistor 319
  - 6.5.1 Output Characteristics 319
  - 6.5.2 Transfer Characteristics 322
  - 6.5.3 Mobility Models 325
  - 6.5.4 Short Channel MOSFET  $I$ – $V$  Characteristics 327
  - 6.5.5 Control of Threshold Voltage 329
  - 6.5.6 Substrate Bias Effects—the “body” effect 332
  - 6.5.7 Subthreshold Characteristics 336
  - 6.5.8 Equivalent Circuit for the MOSFET 338
  - 6.5.9 MOSFET Scaling and Hot Electron Effects 341
  - 6.5.10 Drain-Induced Barrier Lowering 345
  - 6.5.11 Short Channel Effect and Narrow Width Effect 347
  - 6.5.12 Gate-Induced Drain Leakage 349
- 6.6 Advanced MOSFET Structures 350
  - 6.6.1 Metal Gate-High- $k$  350
  - 6.6.2 Enhanced Channel Mobility Materials and Strained Si FETs 351
  - 6.6.3 SOI MOSFETs and FinFETs 353

# 7

## BIPOLAR JUNCTION TRANSISTORS 368

- 7.1 Fundamentals of BJT Operation 368
- 7.2 Amplification with BJTs 372
- 7.3 BJT Fabrication 375
- 7.4 Minority Carrier Distributions and Terminal Currents 378
  - 7.4.1 Solution of the Diffusion Equation in the Base Region 379
  - 7.4.2 Evaluation of the Terminal Currents 381
  - 7.4.3 Approximations of the Terminal Currents 384
  - 7.4.4 Current Transfer Ratio 386
- 7.5 Generalized Biasing 387
  - 7.5.1 The Coupled-Diode Model 388
  - 7.5.2 Charge Control Analysis 393
- 7.6 Switching 395
  - 7.6.1 Cutoff 396
  - 7.6.2 Saturation 397
  - 7.6.3 The Switching Cycle 398
  - 7.6.4 Specifications for Switching Transistors 399
- 7.7 Other Important Effects 400
  - 7.7.1 Drift in the Base Region 401
  - 7.7.2 Base Narrowing 402
  - 7.7.3 Avalanche Breakdown 403
  - 7.7.4 Injection Level; Thermal Effects 405
  - 7.7.5 Base Resistance and Emitter Crowding 406
  - 7.7.6 Gummel–Poon Model 408
  - 7.7.7 Kirk Effect 411
- 7.8 Frequency Limitations of Transistors 414
  - 7.8.1 Capacitance and Charging Times 414
  - 7.8.2 Transit Time Effects 417
  - 7.8.3 Webster Effect 418
  - 7.8.4 High-Frequency Transistors 418
- 7.9 Heterojunction Bipolar Transistors 420

# 8

## OPTOELECTRONIC DEVICES 430

- 8.1 Photodiodes 430
  - 8.1.1 Current and Voltage in an Illuminated Junction 431
  - 8.1.2 Solar Cells 434
  - 8.1.3 Photodetectors 437
  - 8.1.4 Gain, Bandwidth, and Signal-to-Noise Ratio of Photodetectors 439

- 8.2 Light-Emitting Diodes 442
  - 8.2.1 Light-Emitting Materials 443
  - 8.2.2 Fiber-Optic Communications 447
- 8.3 Lasers 450
- 8.4 Semiconductor Lasers 454
  - 8.4.1 Population Inversion at a Junction 455
  - 8.4.2 Emission Spectra for p-n Junction Lasers 457
  - 8.4.3 The Basic Semiconductor Laser 458
  - 8.4.4 Heterojunction Lasers 459
  - 8.4.5 Materials for Semiconductor Lasers 462
  - 8.4.6 Quantum Cascade Lasers 464

## 9

### INTEGRATED CIRCUITS 472

- 9.1 Background 473
  - 9.1.1 Advantages of Integration 473
  - 9.1.2 Types of Integrated Circuits 475
- 9.2 Evolution of Integrated Circuits 476
- 9.3 Monolithic Device Elements 479
  - 9.3.1 CMOS Process Integration 479
  - 9.3.2 Integration of Other Circuit Elements 494
- 9.4 Charge Transfer Devices 500
  - 9.4.1 Dynamic Effects in MOS Capacitors 501
  - 9.4.2 The Basic CCD 502
  - 9.4.3 Improvements on the Basic Structure 503
  - 9.4.4 Applications of CCDs 504
- 9.5 Ultra Large-Scale Integration (ULSI) 505
  - 9.5.1 Logic Devices 507
  - 9.5.2 Semiconductor Memories 517
- 9.6 Testing, Bonding, and Packaging 530
  - 9.6.1 Testing 531
  - 9.6.2 Wire Bonding 531
  - 9.6.3 Flip-Chip Techniques 535
  - 9.6.4 Packaging 535

## 10

### HIGH-FREQUENCY, HIGH-POWER AND NANO-ELECTRONIC DEVICES 541

- 10.1 Tunnel Diodes 541
  - 10.1.1 Degenerate Semiconductors 541
- 10.2 The IMPATT Diode 545
- 10.3 The Gunn Diode 548
  - 10.3.1 The Transferred-Electron Mechanism 548
  - 10.3.2 Formation and Drift of Space Charge Domains 551

10.4 The p-n-p-n Diode	553
10.4.1 Basic Structure	553
10.4.2 The Two-Transistor Analogy	554
10.4.3 Variation of $\alpha$ with Injection	555
10.4.4 Forward-Blocking State	556
10.4.5 Conducting State	557
10.4.6 Triggering Mechanisms	558
10.5 The Semiconductor-Controlled Rectifier	559
10.5.1 Turning off the SCR	560
10.6 Insulated-Gate Bipolar Transistor	561
10.7 Nanoelectronic Devices	564
10.7.1 Zero-Dimensional Quantum Dots	564
10.7.2 One-Dimensional Quantum Wires	566
10.7.3 Two-Dimensional Layered Crystals	567
10.7.4 Spintronic Memory	568
10.7.5 Nanoelectronic Resistive Memory	570

## APPENDICES

I. Definitions of Commonly Used Symbols	575
II. Physical Constants and Conversion Factors	579
III. Properties of Semiconductor Materials	580
IV. Derivation of the Density of States in the Conduction Band	581
V. Derivation of Fermi–Dirac Statistics	586
VI. Dry and Wet Thermal Oxide Thickness Grown on Si (100) as a Function of Time and Temperature	589
VII. Solid Solubilities of Impurities in Si	591
VIII. Diffusivities of Dopants in Si and SiO <sub>2</sub>	592
IX. Projected Range and Straggle as Function of Implant Energy in Si	594

## ANSWERS TO SELECTED SELF QUIZ QUESTIONS 596

## INDEX 600





This book is an introduction to semiconductor devices for undergraduate electrical engineers, other interested students, and practicing engineers and scientists whose understanding of modern electronics needs updating. The book is organized to bring students with a background in sophomore physics to a level of understanding that will allow them to read much of the current literature on new devices and applications.

An undergraduate course in electronic devices has two basic purposes: (1) to provide students with a sound understanding of existing devices, so that their studies of electronic circuits and systems will be meaningful and (2) to develop the basic tools with which they can later learn about newly developed devices and applications. Perhaps the second of these objectives is the more important in the long run; it is clear that engineers and scientists who deal with electronics will continually be called upon to learn about new devices and processes in the future. For this reason, we have tried to incorporate the basics of semiconductor materials and conduction processes in solids, which arise repeatedly in the literature when new devices are explained. Some of these concepts are often omitted in introductory courses, with the view that they are unnecessary for understanding the fundamentals of junctions and transistors. We believe this view neglects the important goal of equipping students for the task of understanding a new device by reading the current literature. Therefore, in this text most of the commonly used semiconductor terms and concepts are introduced and related to a broad range of devices.

---

## GOALS

- updated discussion of MOS devices, both in the underlying theory of ballistic FETs as well as discussion of advanced MOSFETs such as FinFETs, strained Si devices, metal gate/ high-k devices, III-V high channel mobility devices
- updated treatment of optoelectronic devices, including high bandgap nitride semiconductors and quantum cascade lasers
- brand new section on nanoelectronics to introduce students to exciting concepts such as 2D materials including graphene and topological insulators, 1D nanowires and nanotubes, and 0D quantum dots;
- discussion of spintronics, and novel resistive and phase change memories
- about 100 new problems, and current references which extend concepts in the text.

---

## WHAT IS NEW IN THIS EDITION

---

**READING LISTS** As a further aid in developing techniques for independent study, the reading list at the end of each chapter includes a few articles which students can read comfortably as they study this book. We do not expect that students will read all articles recommended in the reading lists; nevertheless, some exposure to periodicals is useful in laying the foundation for a career of constant updating and self-education. We have also added a summary of the key concepts at the end of each chapter.

---

**PROBLEMS** One of the keys to success in understanding this material is to work problems that exercise the concepts. The problems at the end of each chapter are designed to facilitate learning the material. Very few are simple “plug-in” problems. Instead, they are chosen to reinforce or extend the material presented in the chapter. In addition, we have added “self quiz” problems that test the conceptual understanding on the part of the students.

---

**UNITS** In keeping with the goals described above, examples and problems are stated in terms of units commonly used in the semiconductor literature. The basic system of units is rationalized MKS, although cm is often used as a convenient unit of length. Similarly, electron volts (eV) are often used rather than joules (J) to measure the energy of electrons. Units for various quantities are given in Appendices I and II.

---

**PRESENTATION** In presenting this material at the undergraduate level, one must anticipate a few instances which call for a phrase such as “It can be shown . . .” This is always disappointing; on the other hand, the alternative is to delay study of solid state devices until the graduate level, where statistical mechanics, quantum theory, and other advanced background can be freely invoked. Such a delay would result in a more elegant treatment of certain subjects, but it would prevent undergraduate students from enjoying the study of some very exciting devices.

The discussion includes both silicon and compound semiconductors, to reflect the continuing growth in importance for compounds in optoelectronic and high-speed device applications. Topics such as heterojunctions, lattice-matching using ternary and quaternary alloys, variation of band gap with alloy composition, and properties of quantum wells add to the breadth of the discussion. Not to be outdone by the compounds, silicon-based devices have continued their dramatic record of advancement. The discussion of FET structures and Si integrated circuits reflects these advancements. Our objective is not to cover all the latest devices, which can only be done in the journal and conference literature. Instead, we have chosen devices to discuss which are broadly illustrative of important principles.

The first four chapters of the book provide background on the nature of semiconductors and conduction processes in solids (Chapters 3, 4). Included is a brief introduction to quantum concepts (Chapter 2) for those students who do not already have this background from other courses. Chapter 5 describes the p-n junction and some of its applications. Chapters 6 and 7 deal with the principles of transistor operation. Chapter 8 covers optoelectronics, and Chapter 9 discusses integrated circuits. Chapter 10 applies the theory of junctions and conduction processes to microwave and power devices. A completely new section on nanoelectronics has been added. All of the devices covered are important in today's electronics; furthermore, learning about these devices should be an enjoyable and rewarding experience. We hope this book provides that kind of experience for its readers.

The seventh edition benefits greatly from comments and suggestions provided by students and teachers of the first six editions. The book's readers have generously provided comments which have been invaluable in developing the present version. We remain indebted to those persons mentioned in the Preface of the first six editions, who contributed so much to the development of the book. In particular, Nick Holonyak has been a source of continuing information and inspiration for all seven editions. Additional thanks go to our colleagues at UT–Austin who have provided special assistance, particularly Leonard Frank Register, Emanuel Tutuc, Ray Chen, Ananth Dodabalapur, Seth Bank, Misha Belkin, Zheng Wang, Neal Hall, Deji Akinwande, Jack Lee, and Dean Neikirk. Hema Movva provided useful assistance with the typing of the homework solutions. We thank the many companies and organizations cited in the figure captions for generously providing photographs and illustrations of devices and fabrication processes. Bob Doering at TI, Mark Bohr at Intel, Chandra Mouli at Micron, Babu Chalamala at MEMC and Kevin Lally at TEL deserve special mention for the new pictures in this edition. Finally, we recall with gratitude many years of association with Joe Campbell, Karl Hess, and the late Al Tasch, valued colleagues and friends.

#### ACKNOWLEDGMENTS

*Ben G. Streetman*  
*Sanjay Kumar Banerjee*

The publishers would like to thank the following for their contribution to the Global Edition:

#### Contributor:

*Rikmantra Basu*, Assistant Professor, Electronics and Communication Engineering (ECE) Department, National Institute of Technology (NIT) Delhi

#### Reviewers:

*Sunanda Khosla* (writer)

*Prof. Tan Chuan Seng*, Nanyang Technological University

*Prof. Dr. habil. Jörg Schulze*, University of Stuttgart



PRENTICE HALL SERIES  
IN SOLID STATE PHYSICAL ELECTRONICS

Nick Holonyak Jr., Editor

Cheo FIBER OPTICS: DEVICES AND SYSTEMS SECOND EDITION

Haus WAVES AND FIELDS IN OPTOELECTRONICS

Kroemer QUANTUM MECHANICS FOR ENGINEERING, MATERIALS SCIENCE,  
AND APPLIED PHYSICS

Nussbaum CONTEMPORARY OPTICS FOR SCIENTISTS AND ENGINEERS

Peyghambarian/Koch/Mysyrowicz INTRODUCTION TO SEMICONDUCTOR OPTICS

Shur PHYSICS OF SEMICONDUCTOR DEVICES

Soclof DESIGN AND APPLICATIONS OF ANALOG INTEGRATED CIRCUITS

Streetman/Banerjee SOLID STATE ELECTRONIC DEVICES SEVENTH EDITION

Verdeyen LASER ELECTRONICS THIRD EDITION

Wolfe/Holonyak/Stillman PHYSICAL PROPERTIES OF SEMICONDUCTORS

## ABOUT THE AUTHORS



*Ben G. Streetman* is Dean Emeritus of the Cockrell School of Engineering at The University of Texas at Austin. He is an Emeritus Professor of Electrical and Computer Engineering, where he held the Dula D. Cockrell Centennial Chair. He was the founding Director of the Microelectronics Research Center (1984–1996) and served as Dean of Engineering from 1996 to 2008. His teaching and research interests involve semiconductor materials and devices. After receiving a Ph.D. from The University of Texas at Austin (1966) he was on the faculty (1966–1982) of the University of Illinois at Urbana-Champaign. He returned to The University of Texas at Austin in 1982. His honors include the Education Medal of the Institute of Electrical and Electronics Engineers (IEEE), the Frederick Emmons Terman Medal of the American Society for Engineering Education (ASEE), and the Heinrich Welker Medal from the International Conference on Compound Semiconductors. He is a member of the National Academy of Engineering and the American Academy of Arts and Sciences. He is a Fellow of the IEEE and the Electrochemical Society. He has been honored as a Distinguished Alumnus of The University of Texas at Austin and as a Distinguished Graduate of the UT College of Engineering. He has received the General Dynamics Award for Excellence in Engineering Teaching, and was honored by the Parents' Association as a Teaching Fellow for outstanding teaching of undergraduates. He has served on numerous panels and committees in industry and government, and several corporate boards. He has published more than 290 articles in the technical literature. Thirty-four students of Electrical Engineering, Materials Science, and Physics have received their Ph.D.s under his direction.



*Sanjay Kumar Banerjee* is the Cockrell Chair Professor of Electrical and Computer Engineering, and Director of the Microelectronics Research Center at The University of Texas at Austin. He received his B.Tech. from the Indian Institute of Technology, Kharagpur, and his M.S. and Ph.D. from the University of Illinois at Urbana-Champaign in 1979, 1981, and 1983, respectively, in electrical engineering. He worked at TI from 1983–1987 on the world’s first 4Megabit DRAM, for which he was a co-recipient of an ISSCC Best Paper Award. He has more than 900 archival refereed publications and conference papers, 30 U.S. patents, and has supervised over 50 Ph.D. students. His honors include the NSF Presidential Young Investigator Award (1988), the Texas Atomic Energy Centennial Fellowship (1990–1997), Cullen Professorship (1997–2001), and the Hocott Research Award from the University of Texas. He has received the ECS Callinan Award (2003), Industrial R&D 100 Award (2004), Distinguished Alumnus Award, IIT (2005), IEEE Millennium Medal (2000) and IEEE Andrew S. Grove Award (2014). He is a Fellow of IEEE, APS and AAAS. He is interested in beyond-CMOS nanoelectronic transistors based on 2D materials and spintronics, fabrication and modeling of advanced MOSFETs, and solar cells.

---

# Solid State Electronic Devices





---

## Chapter 1

# Crystal Properties and Growth of Semiconductors

---

### OBJECTIVES

1. Describe what a semiconductor is
2. Perform simple calculations about crystals
3. Understand what is involved in bulk Czochralski and thin-film epitaxial crystal growth
4. Learn about crystal defects

In studying solid state electronic devices we are interested primarily in the electrical behavior of solids. However, we shall see in later chapters that the transport of charge through a metal or a semiconductor depends not only on the properties of the electron but also on the arrangement of atoms in the solid. In this chapter we shall discuss some of the physical properties of semiconductors compared with other solids, the atomic arrangements of various materials, and some methods of growing semiconductor crystals. Topics such as crystal structure and crystal growth technology are often the subjects of books rather than introductory chapters; thus we shall consider only a few of the more important and fundamental ideas that form the basis for understanding electronic properties of semiconductors and device fabrication.

Semiconductors are a group of materials having electrical conductivities intermediate between metals and insulators. It is significant that the conductivity of these materials can be varied over orders of magnitude by changes in temperature, optical excitation, and impurity content. This variability of electrical properties makes the semiconductor materials natural choices for electronic device investigations.

Semiconductor materials are found in column IV and neighboring columns of the periodic table (Table 1–1). The column IV semiconductors, silicon and germanium, are called *elemental* semiconductors because they are composed of single species of atoms. In addition to the elemental materials, compounds of column III and column V atoms, as well as certain combinations from II and VI, and from IV, make up the *compound* semiconductors.

### 1.1 SEMICONDUCTOR MATERIALS

As Table 1–1 indicates, there are numerous semiconductor materials. As we shall see, the wide variety of electronic and optical properties of these semiconductors provides the device engineer with great flexibility in the design of electronic and optoelectronic functions. The elemental semiconductor Ge was widely used in the early days of semiconductor development for transistors and diodes. Silicon is now used for the majority of rectifiers, transistors, and integrated circuits (ICs). However, the compounds are widely used in high-speed devices and devices requiring the emission or absorption of light. The two-element (*binary*) III–V compounds such as GaN, GaP, and GaAs are common in light-emitting diodes (LEDs). As discussed in Section 1.2.4, three-element (*ternary*) compounds such as GaAsP and four-element (*quaternary*) compounds such as InGaAsP can be grown to provide added flexibility in choosing materials properties.

Fluorescent materials such as those used in television screens usually are II–VI compound semiconductors such as ZnS. Light detectors are commonly made with InSb, CdSe, or other compounds such as PbTe and HgCdTe. Si and Ge are also widely used as infrared and nuclear radiation detectors. Light-emitting diodes are made using GaN and other III–V compounds. Semiconductor lasers are made using GaAs, AlGaAs, and other ternary and quaternary compounds.

One of the most important characteristics of a semiconductor, which distinguishes it from metals and insulators, is its *energy band gap*. This property, which we will discuss in detail in Chapter 3, determines among other things the wavelengths of light that can be absorbed or emitted by the semiconductor. For example, the band gap of GaAs is about 1.43 electron volts (eV), which

**Table 1–1** Common semiconductor materials: (a) the portion of the periodic table where semiconductors occur; (b) elemental and compound semiconductors.

(a)	II	III	IV	V	VI
		B	C	N	
		Al	Si	P	S
	Zn	Ga	Ge	As	Se
	Cd	In		Sb	Te
(b)	Elemental	IV compounds	Binary III–V compounds	Binary II–VI compounds	
	Si	SiC	AlP	ZnS	
	Ge	SiGe	AlAs	ZnSe	
			AlSb	ZnTe	
			GaN	CdS	
			GaP	CdSe	
			GaAs	CdTe	
			GaSb		
			InP		
			InAs		
			InSb		

corresponds to light wavelengths in the near infrared. In contrast, GaP has a band gap of about 2.3 eV, corresponding to wavelengths in the green portion of the spectrum.<sup>1</sup> The band gap  $E_g$  for various semiconductor materials is listed along with other properties in Appendix III. As a result of the wide variety of semiconductor band gaps, LEDs and lasers can be constructed with wavelengths over a broad range of the infrared and visible portions of the spectrum.

The electronic and optical properties of semiconductor materials are strongly affected by impurities, which may be added in precisely controlled amounts. Such impurities are used to vary the conductivities of semiconductors over wide ranges and even to alter the nature of the conduction processes from conduction by negative charge carriers to positive charge carriers. For example, an impurity concentration of one part per million can change a sample of Si from a poor conductor to a good conductor of electric current. This process of controlled addition of impurities, called *doping*, will be discussed in detail in subsequent chapters.

To investigate these useful properties of semiconductors, it is necessary to understand the atomic arrangements in the materials. Obviously, if slight alterations in purity of the original material can produce such dramatic changes in electrical properties, then the nature and specific arrangement of atoms in each semiconductor must be of critical importance. Therefore, we begin our study of semiconductors with a brief introduction to crystal structure.

In this section we discuss the arrangements of atoms in various solids. We shall distinguish between single crystals and other forms of materials and then investigate the periodicity of crystal lattices. Certain important crystallographic terms will be defined and illustrated in reference to crystals having a basic cubic structure. These definitions will allow us to refer to certain planes and directions within a lattice. Finally, we shall investigate the diamond lattice; this structure, with some variations, is typical of most of the semiconductor materials used in electronic devices.

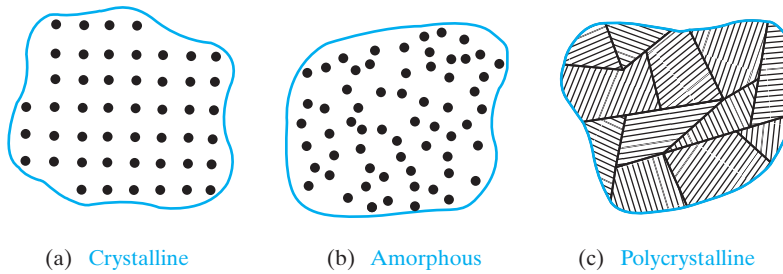
## 1.2 CRYSTAL LATTICES

### 1.2.1 Periodic Structures

A crystalline solid is distinguished by the fact that the atoms making up the crystal are arranged in a periodic fashion. That is, there is some basic arrangement of atoms that is repeated throughout the entire solid. Thus the crystal appears exactly the same at one point as it does at a series of other equivalent points, once the basic periodicity is discovered. However, not all solids are crystals (Fig. 1-1); some have no periodic structure at all (*amorphous* solids), and others are composed of many small regions of single-crystal material (*polycrystalline* solids). The high-resolution micrograph shown in Fig. 6-33 illustrates the periodic array of atoms in the single-crystal silicon of a transistor channel compared with the amorphous SiO<sub>2</sub> (glass) of the oxide layer.

<sup>1</sup>The conversion between the energy  $E$  of a photon of light (eV) and its wavelength  $\lambda$  ( $\mu\text{m}$ ) is  $\lambda = 1.24/E$ . For GaAs,  $\lambda = 1.24/1.43 = 0.87 \mu\text{m}$ .



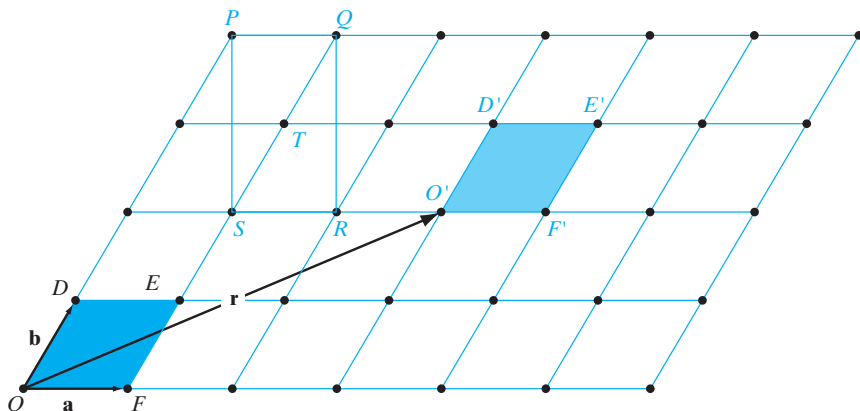
**Figure 1-1**

Three types of solids, classified according to atomic arrangement: (a) crystalline and (b) amorphous materials are illustrated by microscopic views of the atoms, whereas (c) polycrystalline structure is illustrated by a more macroscopic view of adjacent single-crystalline regions, such as (a).

The periodicity in a crystal is defined in terms of a symmetric array of points in space called the *lattice*. We can add atoms at each lattice point in an arrangement called a *basis*, which can be one atom or a group of atoms having the same spatial arrangement, to get a *crystal*. In every case, the lattice contains a volume or *cell* that represents the entire lattice and is regularly repeated throughout the crystal. As an example of such a lattice, Fig. 1-2 shows a two-dimensional arrangement of atoms called a rhombic lattice, with a *primitive* cell ODEF, which is the smallest such cell. Notice that we can define vectors **a** and **b** such that if the primitive cell is translated by integral multiples of these vectors, a new primitive cell identical to the original is found (e.g., O'D'E'F'). These vectors, **a** and **b** (and **c** if the lattice is three dimensional), are called the *primitive vectors* for the lattice. Points within the lattice are indistinguishable if the vector between the points is

$$\mathbf{r} = p\mathbf{a} + q\mathbf{b} + s\mathbf{c} \quad (1-1)$$

where  $p$ ,  $q$ , and  $s$  are integers. The primitive cell shown has lattice points *only* at the corners of the cell. The primitive cell is not unique, but it must cover



**Figure 1-2**  
A two-dimensional  
lattice showing  
translation of  
a unit cell by  
 $\mathbf{r} = 3\mathbf{a} + 2\mathbf{b}$ .

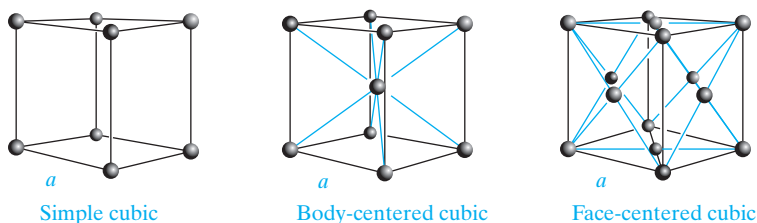
the entire volume of the crystal (without missing or extra bits) by translations by integer numbers of primitive vectors, and it can have *only* one lattice point per cell. The convention is to choose the smallest primitive vectors. Note that, in the primitive cell shown in Fig. 1–2, the lattice points at the corners are shared with adjacent cells; thus, the *effective* number of lattice points belonging to the primitive cell is unity. Since there are many different ways of placing atoms in a volume, the distances and orientation between atoms can take many forms, leading to different lattice and crystal structures. It is important to remember that the *symmetry* determines the lattice, not the magnitudes of the distances between the lattice points.

In many lattices, however, the primitive cell is not the most convenient to work with. For example, in Fig. 1–2, we see that the rhombic arrangement of the lattice points is such that it can also be considered to be rectangular (PQRS) with a lattice point in the center at T (a so-called *centered rectangular* lattice). (Note that this is not true of all rhombic lattices!) Clearly, it is simpler to deal with a rectangle rather than a rhombus. So, in this case we can choose to work with a larger rectangular *unit cell*, PQRS, rather than the smallest primitive cell, ODEF. A unit cell allows lattice points not only at the corners but also at the face center (and body center in 3-D) if necessary. It is sometimes used instead of the primitive cell if it can represent the symmetry of the lattice better (in this example “centered rectangular” two-dimensional lattice). It replicates the lattice by integer translations of *basis* vectors.

The importance of the unit cell lies in the fact that we can analyze the crystal as a whole by investigating a representative volume. For example, from the unit cell we can find the distances between nearest atoms and next nearest atoms for calculation of the forces holding the lattice together; we can look at the fraction of the unit cell volume filled by atoms and relate the density of the solid to the atomic arrangement. But even more important for our interest in electronic devices, the properties of the periodic crystal lattice determine the allowed energies of electrons that participate in the conduction process. Thus the lattice determines not only the mechanical properties of the crystal but also its electrical properties.

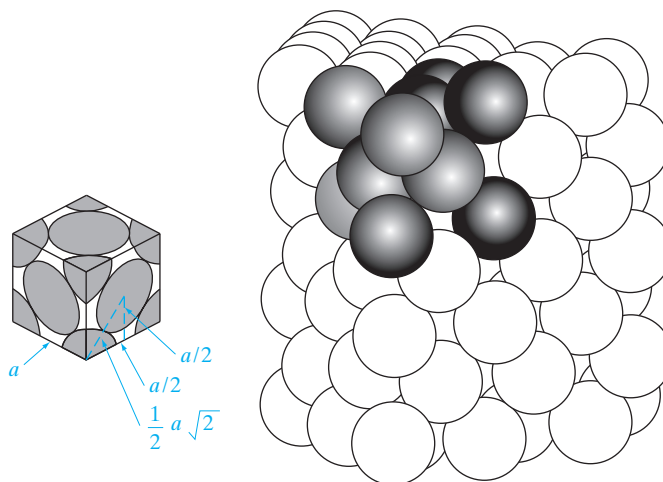
### 1.2.2 Cubic Lattices

The simplest three-dimensional lattice is one in which the unit cell is a cubic volume, such as the three cells shown in Fig. 1–3. The *simple cubic* structure (abbreviated *sc*) has an atom located at each corner of the unit cell. The



**Figure 1–3**  
Unit cells for three types of cubic lattice structures.

**Figure 1–4**  
Packing of hard  
spheres in an fcc  
lattice.



*body-centered cubic (bcc)* lattice has an additional atom at the center of the cube, and the *face-centered cubic (fcc)* unit cell has atoms at the eight corners and centered on the six faces. All three structures have different primitive cells, but the same cubic unit cell. We will generally work with unit cells.

As atoms are packed into the lattice in any of these arrangements, the distances between neighboring atoms will be determined by a balance between the forces that attract them together and other forces that hold them apart. We shall discuss the nature of these forces for particular solids in Section 3.1.1. For now, we can calculate the maximum fraction of the lattice volume that can be filled with atoms by approximating the atoms as hard spheres. For example, Fig. 1–4 illustrates the packing of spheres in a fcc cell of side  $a$ , such that the nearest neighbors touch. The dimension  $a$  for a cubic unit cell is called the *lattice constant*. For the fcc lattice the nearest neighbor distance is one-half the diagonal of a face, or  $\frac{1}{2}(a\sqrt{2})$ . Therefore, for the atom centered on the face to just touch the atoms at each corner of the face, the radius of the sphere must be one-half the nearest neighbor distance, or  $\frac{1}{4}(a\sqrt{2})$ .

**EXAMPLE 1–1** Find the fraction of the fcc unit cell volume filled with hard spheres.

**SOLUTION**

$$\text{Nearest atom separation} = \frac{5\sqrt{2}}{2} \text{ \AA} = 3.54 \text{ \AA}$$

$$\text{Tetrahedral radius} = 1.77 \text{ \AA}$$

$$\text{Volume of each atom} = 23.14 \text{ \AA}^3$$

$$\text{Number of atoms per cube} = 6 \cdot \frac{1}{2} + 8 \cdot \frac{1}{8} = 4 \text{ atoms}$$

$$\text{Packing fraction} = \frac{23.1 \text{ \AA}^3 \cdot 4}{(5 \text{ \AA})^3} = 0.74 = 74\%$$

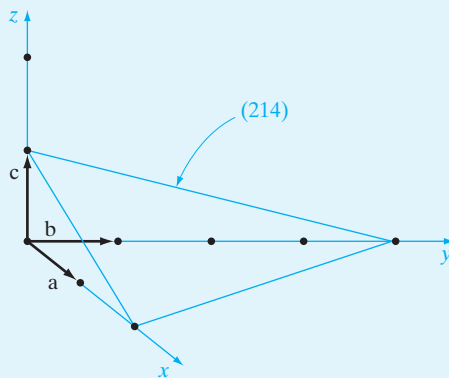
### 1.2.3 Planes and Directions

In discussing crystals it is very helpful to be able to refer to planes and directions within the lattice. The notation system generally adopted uses a set of three integers to describe the position of a plane or the direction of a vector within the lattice. We first set up an  $xyz$  coordinate system with the origin at any lattice point (it does not matter which one because they are all equivalent!), and the axes are lined up with the edges of the cubic unit cell. The three integers describing a particular plane are found in the following way:

1. Find the intercepts of the plane with the crystal axes and express these intercepts as integral multiples of the basis vectors (the plane can be moved in and out from the origin, retaining its orientation, until such an integral intercept is discovered on each axis).
2. Take the reciprocals of the three integers found in step 1 and reduce these to the smallest set of integers  $h$ ,  $k$ , and  $l$ , which have the same relationship to each other as the three reciprocals.
3. Label the plane ( $hkl$ ).

The plane illustrated in Fig. 1–5 has intercepts at  $2a$ ,  $4b$ , and  $1c$  along the three crystal axes. Taking the reciprocals of these intercepts, we get  $\frac{1}{2}$ ,  $\frac{1}{4}$ , and  $1$ . These three fractions have the same relationship to each other as the integers 2, 1, and 4 (obtained by multiplying each fraction by 4). Thus the plane can be referred to as a (214) plane. The only exception is if the intercept is a fraction of the lattice constant  $a$ . In that case, we do not reduce it to the lowest set of integers. For example, in Fig. 1–3, planes parallel to the cube faces, but going through the body center atoms in the bcc lattice, would be (200) and not (100).

#### EXAMPLE 1–2



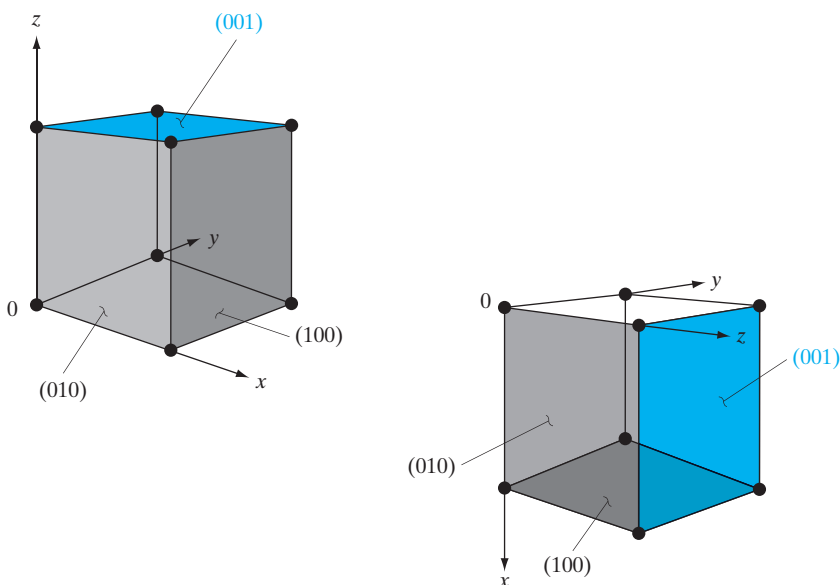
**Figure 1–5**  
A (214) crystal plane.

The three integers  $h$ ,  $k$ , and  $l$  are called the *Miller indices*; these three numbers define a set of parallel planes in the lattice. One advantage of taking the reciprocals of the intercepts is avoidance of infinities in the notation. One

intercept is infinity for a plane parallel to an axis; however, the reciprocal of such an intercept is taken as zero. If a plane contains one of the axes, it is parallel to that axis and has a zero reciprocal intercept. If a plane passes through the origin, it can be translated to a parallel position for calculation of the Miller indices. If an intercept occurs on the negative branch of an axis, the minus sign is placed above the Miller index for convenience, such as  $(h\bar{k}l)$ .

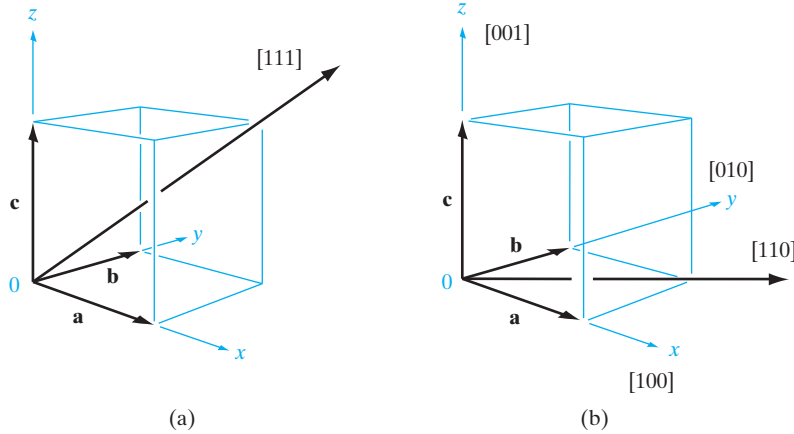
From a crystallographic point of view, many planes in a lattice are equivalent; that is, a plane with given Miller indices can be shifted about in the lattice simply by choice of the position and orientation of the unit cell. The indices of such equivalent planes are enclosed in braces  $\{ \}$  instead of parentheses. For example, in the cubic lattice of Fig. 1–6, all the cube faces are crystallographically equivalent in that the unit cell can be rotated in various directions and still appear the same. The six equivalent faces are collectively designated as  $\{100\}$ .

A direction in a lattice is expressed as a set of three integers with the same relationship as the components of a vector in that direction. The three vector components are expressed in multiples of the basis vectors, and the three integers are reduced to their smallest values while retaining the relationship among them. For example, the body diagonal in the cubic lattice (Fig. 1–7a) is composed of the components  $1\mathbf{a}$ ,  $1\mathbf{b}$ , and  $1\mathbf{c}$ ; therefore, this diagonal is the  $[111]$  direction. (Brackets are used for direction indices.) As in the case of planes, many directions in a lattice are equivalent, depending only on the arbitrary choice of orientation for the axes. Such equivalent direction indices are placed in angular brackets  $\langle \rangle$ . For example, the crystal axes in the cubic lattice  $[100]$ ,  $[010]$ , and  $[001]$  are all equivalent and are called  $\langle 100 \rangle$  directions (Fig. 1–7b).



**Figure 1–6**  
Equivalence of  
the cube faces  
( $\{100\}$  planes)  
by rotation of the  
unit cell within the  
cubic lattice.





**Figure 1-7**  
Crystal directions  
in the cubic  
lattice.

Two useful relationships in terms of Miller indices describe the distance between planes and angles between directions. The distance  $d$  between two adjacent planes labeled  $(hkl)$  is given in terms of the lattice constant,  $a$ , as

$$d = a/(h^2 + k^2 + l^2)^{1/2} \quad (1-2a)$$

The angle  $\theta$  between two different Miller index directions is given by

$$\cos \theta = \{h_1 h_2 + k_1 k_2 + l_1 l_2\} \{ (h_1^2 + k_1^2 + l_1^2)^{1/2} (h_2^2 + k_2^2 + l_2^2)^{1/2} \} \quad (1-2b)$$

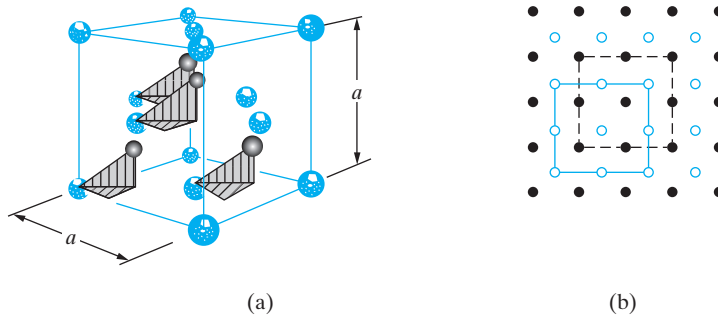
Comparing Figs. 1-6 and 1-7, we notice that in cubic lattices a direction  $[hkl]$  is perpendicular to the plane  $(hkl)$ . This is convenient in analyzing lattices with cubic unit cells, but it should be remembered that it is not necessarily true in noncubic systems.

### 1.2.4 The Diamond Lattice

The basic crystal structure for many important semiconductors is the fcc lattice with a basis of two atoms, giving rise to the *diamond* structure, characteristic of Si, Ge, and C in the diamond form. In many compound semiconductors, atoms are arranged in a basic diamond structure, but are different on alternating sites. This is called a *zinc blende* structure and is typical of the III-V compounds. One of the simplest ways of stating the construction of the diamond structure is the following:

The diamond structure can be thought of as an fcc lattice with an extra atom placed at  $\mathbf{a}/4 + \mathbf{b}/4 + \mathbf{c}/4$  from each of the fcc atoms.

Figure 1-8a illustrates the construction of a diamond lattice from an fcc unit cell. We notice that when the vectors are drawn with components one-fourth of the cube edge in each direction, only four additional points within the same unit cell are reached. Vectors drawn from any of the other fcc atoms simply determine corresponding points in adjacent unit cells. This method of

**Figure 1-8**

Diamond lattice structure: (a) a unit cell of the diamond lattice constructed by placing atoms  $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}$  from each atom in an fcc; (b) top view (along any  $\langle 100 \rangle$  direction) of an extended diamond lattice. The colored circles indicate one fcc sublattice and the black circles indicate the interpenetrating fcc.

constructing the diamond lattice implies that the original fcc has associated with it a second interpenetrating fcc displaced by  $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}$ . The two interpenetrating fcc *sublattices* can be visualized by looking down on the unit cell of Fig. 1-8a from the top (or along any  $\langle 100 \rangle$  direction). In the top view of Fig. 1-8b, atoms belonging to the original fcc are represented by open circles, and the interpenetrating sublattice is shaded. If the atoms are all similar, we call this structure a diamond lattice; if the atoms differ on alternating sites, it is a zinc blende structure. For example, if one fcc sublattice is composed of Ga atoms and the interpenetrating sublattice is As, the zinc blende structure of GaAs results. Most of the compound semiconductors have this type of lattice, although some of the II-VI compounds are arranged in a slightly different structure called the *wurtzite* lattice. We shall restrict our discussion here to the diamond and zinc blende structures, since they are typical of most of the commonly used semiconductors.

**EXAMPLE 1-3**

Calculate the volume density of Si atoms (number of atoms/cm<sup>3</sup>), given that the lattice constant of Si is 5.43 Å. Calculate the areal density of atoms (number/cm<sup>2</sup>) on the (100) plane.

**SOLUTION**

On the (100) plane, we have four atoms on corners and one on the face center.

$$(100) \text{ plane: } \frac{4 \times \frac{1}{4} + 1}{(5.43 \times 10^{-8})(5.43 \times 10^{-8})} = 6.8 \times 10^{14} \text{ cm}^{-2}$$

For Si, we have eight corner lattice points, six face centered points, and two atoms

$$\text{Number of atoms per cube} = \left(8 \times \frac{1}{8} + \frac{1}{2} \times 6\right) \times 2 = 8$$

$$\text{Volume density} = \frac{8}{(5.43 \times 10^{-8})^3} = 5.00 \times 10^{22} \text{ cm}^{-3}$$

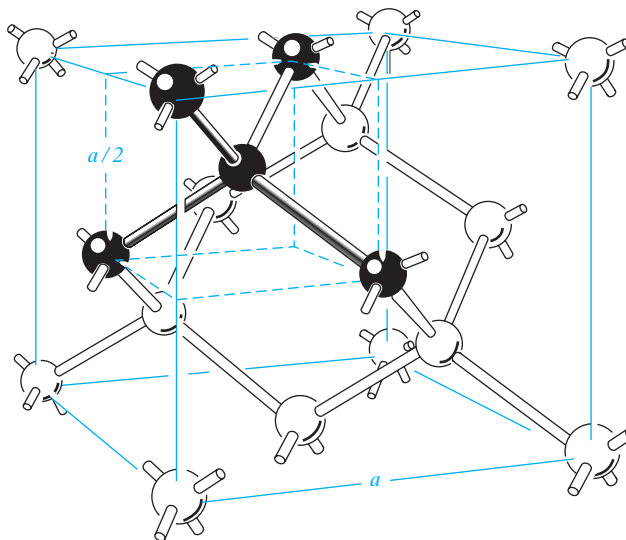
A particularly interesting and useful feature of the III–V compounds is the ability to vary the mixture of elements on each of the two interpenetrating fcc sublattices of the zinc blende crystal. For example, in the ternary compound AlGaAs, it is possible to vary the composition of the ternary alloy by choosing the fraction of Al or Ga atoms on the column III sublattice. It is common to represent the composition by assigning subscripts to the various elements. For example,  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  refers to a ternary alloy in which the column III sublattice in the zinc blende structure contains a fraction  $x$  of Al atoms and  $1-x$  of Ga atoms. The composition  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  has 30 percent Al and 70 percent Ga on the column III sites, with the interpenetrating column V sublattice occupied entirely by As atoms. It is extremely useful to be able to grow ternary alloy crystals such as this with a given composition. For the  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  example we can grow crystals over the entire composition range from  $x = 0$  to  $x = 1$ , thus varying the electronic and optical properties of the material from that of GaAs ( $x = 0$ ) to that of AlAs ( $x = 1$ ). To vary the properties even further, it is possible to grow four-element (quaternary) compounds such as  $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{P}_{1-y}$  having a very wide range of properties.

It is important from an electronic point of view to notice that each atom in the diamond and zinc blende structures is surrounded by four nearest neighbors (Fig. 1–9). The importance of this relationship of each atom to its neighbors will become evident in Section 3.1.1 when we discuss the bonding forces which hold the lattice together.

The fact that atoms in a crystal are arranged in certain planes is important to many of the mechanical, metallurgical, and chemical properties of the material. For example, crystals often can be cleaved along certain atomic planes, resulting in exceptionally planar surfaces. This is a familiar result in cleaved diamonds for jewelry; the facets of a diamond reveal clearly the triangular, hexagonal, and rectangular symmetries of intersecting planes in various crystallographic directions. Semiconductors with diamond and zinc blende lattices have similar cleavage planes. Chemical reactions, such as etching of the crystal, often take place preferentially along certain directions. These properties serve as interesting illustrations of crystal symmetry, but in addition, each plays an important role in fabrication processes for many semiconductor devices.

**Figure 1-9**

Diamond  
lattice unit cell,  
showing the four  
nearest neighbor  
structure. (From  
*Electrons and Holes  
in Semiconductors*  
by W. Shockley,  
© 1950 by Litton  
Educational  
Publishing Co.,  
Inc.; by permission  
of Van Nostrand  
Reinhold Co., Inc.)



### 1.3 BULK CRYSTAL GROWTH

The progress of solid state device technology since the invention of the transistor in 1948 has depended not only on the development of device concepts but also on the improvement of materials. For example, the fact that ICs can be made today is the result of a considerable breakthrough in the growth of pure, single-crystal Si in the early and mid-1950s. The requirements on the growing of device-grade semiconductor crystals are more stringent than those for any other materials. Not only must semiconductors be available in large single crystals, but also the purity must be controlled within extremely close limits. For example, Si crystals now being used in devices are grown with concentrations of most impurities of less than one part in ten billion. Such purities require careful handling and treatment of the material at each step of the manufacturing process.

#### 1.3.1 Starting Materials

The raw feedstock for Si crystal is silicon dioxide ( $\text{SiO}_2$ ). We react  $\text{SiO}_2$  with C in the form of coke in an arc furnace at very high temperatures ( $\sim 1800^\circ\text{C}$ ) to reduce  $\text{SiO}_2$  according to the following reaction:

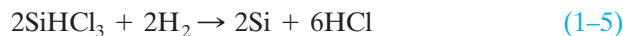


This forms metallurgical grade Si (MGS) which has impurities such as Fe, Al, and heavy metals at levels of several hundred to several thousand parts per million (ppm). Refer back to Example 1-3 to see that 1 ppm of Si corresponds to an impurity level of  $5 \times 10^{16} \text{ cm}^{-3}$ . While MGS is clean enough for metallurgical applications such as using Si to make stainless steel, it is not pure enough for electronic applications; it is also not single crystal.

The MGS is refined further to yield semiconductor-grade or electronic-grade Si (EGS), in which the levels of impurities are reduced to parts per billion or ppb ( $1 \text{ ppb} = 5 \times 10^{13} \text{ cm}^{-3}$ ). This involves reacting the MGS with dry HCl according to the following reaction to form trichlorosilane,  $\text{SiHCl}_3$ , which is a liquid with a boiling point of  $32^\circ\text{C}$ :



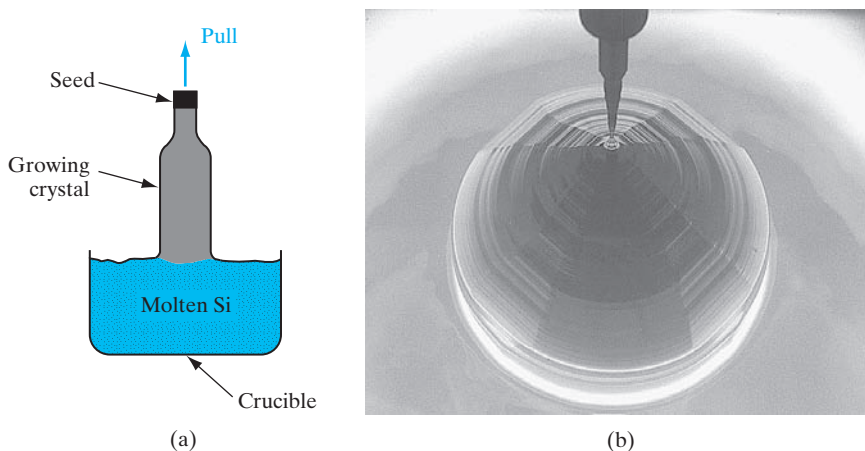
Along with  $\text{SiHCl}_3$ , chlorides of impurities such as  $\text{FeCl}_3$  are formed which fortunately have boiling points that are different from that of  $\text{SiHCl}_3$ . This allows a technique called fractional distillation to be used, in which we heat up the mixture of  $\text{SiHCl}_3$  and the impurity chlorides, and condense the vapors in different distillation towers held at appropriate temperatures. We can thereby separate pure  $\text{SiHCl}_3$  from the impurities.  $\text{SiHCl}_3$  is then converted to highly pure EGS by reaction with  $\text{H}_2$ ,



### 1.3.2 Growth of Single-Crystal Ingots

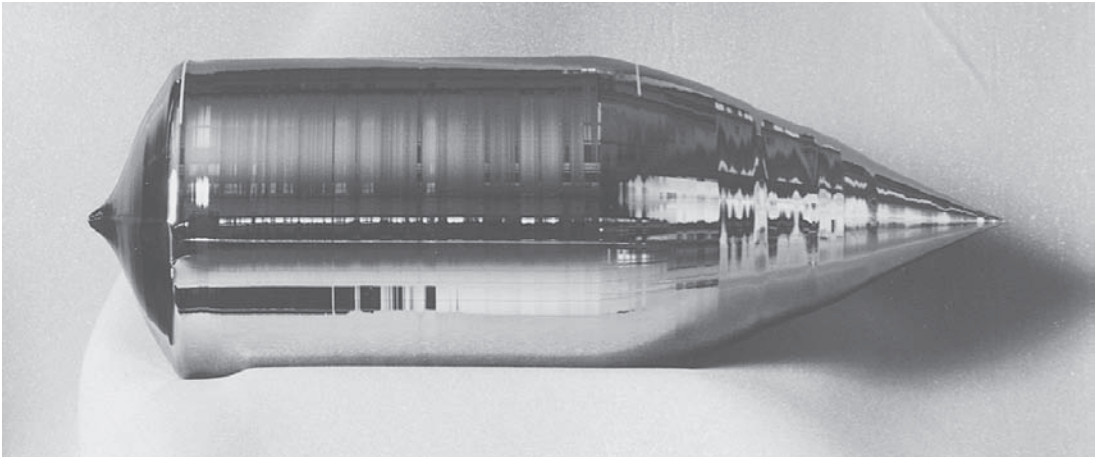
Next, we have to convert the high purity but still polycrystalline EGS to single-crystal Si ingots or boules. This is generally done today by a process commonly called the *Czochralski* method. In order to grow single-crystal material, it is necessary to have a seed crystal which can provide a template for growth. We melt the EGS in a quartz-lined graphite crucible by resistively heating it to the melting point of Si ( $1412^\circ\text{C}$ ).

A seed crystal is lowered into the molten material and then is raised slowly, allowing the crystal to grow onto the seed (Fig. 1–10a). Generally, the



**Figure 1–10**

Pulling of a Si crystal from the melt (Czochralski method): (a) schematic diagram of the crystal growth process; (b) an 8-in. diameter,  $\langle 100 \rangle$  oriented Si crystal being pulled from the melt. (Photograph courtesy of MEMC Electronics Intl.)



(a)



(b)

**Figure 1-11**

(a) Silicon crystal grown by the Czochralski method. This large single-crystal ingot provides 300 mm (12-in.) diameter wafers when sliced using a saw. The ingot is about 1.0 m long (including the tapered regions), and weighs about 140 kg. (b) technician holding a 300 mm wafer. (Photograph courtesy of MEMC Electronics Intl.)

crystal is rotated slowly as it grows to provide a slight stirring of the melt and to average out any temperature variations that would cause inhomogeneous solidification. This technique is widely used in growing Si, Ge, and some of the compound semiconductors.

In pulling compounds such as GaAs from the melt, it is necessary to prevent volatile elements (e.g., As) from vaporizing. In one method a layer of  $B_2O_3$ , which is dense and viscous when molten, floats on the surface of the molten GaAs to prevent As evaporation. This growth method is called *liquid-encapsulated Czochralski (LEC)* growth.

In Czochralski crystal growth, the shape of the ingot is determined by a combination of the tendency of the cross section to assume a polygonal shape due to the crystal structure and the influence of surface tension, which encourages a circular cross section. The crystal facets are noticeable in the initial growth near the seed crystal in Fig. 1–10b. However, the cross section of the large ingot in Fig. 1–11 is almost circular.

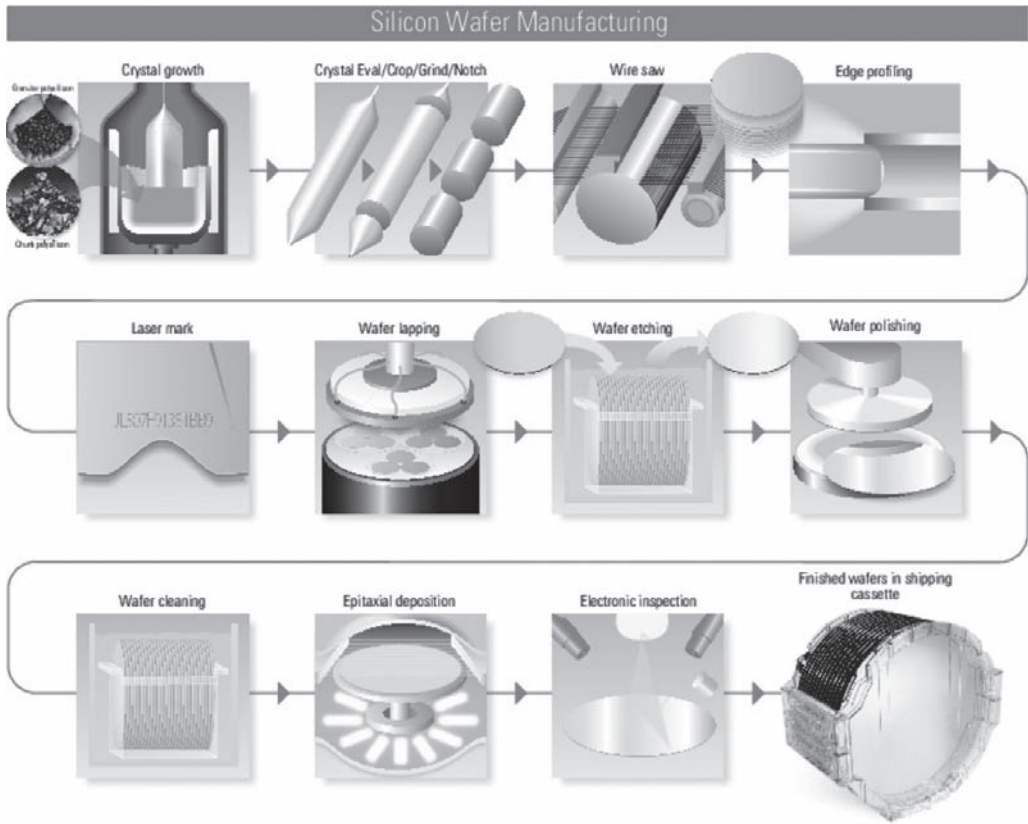
In the fabrication of Si ICs (Chapter 9) it is economical to use very large Si wafers, so that many IC chips can be made simultaneously. As a result, considerable research and development have gone into methods for growing very large Si crystals. For example, Fig. 1–11 illustrates a ~12-inch or 300-mm diameter Si ingot, 1.0 m long and weighing 140 kg, and a 300 mm wafer.

### 1.3.3 Wafers

After the single-crystal ingot is grown, it is then mechanically processed to manufacture wafers. The first step involves mechanically grinding the more-or-less cylindrical ingot into a perfect cylinder with a precisely controlled diameter. This is important because in a modern IC fabrication facility, many processing tools and wafer handling robots require tight tolerances on the size of the wafers. Using X-ray crystallography, crystal planes in the ingot are identified. For reasons discussed in Section 6.4.3, most Si ingots are grown along the  $\langle 100 \rangle$  direction (Fig. 1–10). For such ingots, a small notch is ground on one side of the cylinder to delineate a  $\{110\}$  face of the crystal. This is useful because for  $\langle 100 \rangle$  Si wafers, the  $\{110\}$  cleavage planes are orthogonal to each other. This notch then allows the individual IC chips to be made oriented along  $\{110\}$  planes so that when the chips are sawed apart, there is less chance of spurious cleavage of the crystal, which could cause good chips to be lost.

Next, the Si cylinder is sawed into individual wafers about 775  $\mu\text{m}$  thick, by using a diamond-tipped inner-hole blade saw, or a wire saw (Fig. 1–12). State-of-the-art wafers today are 300 mm in diameter; the next size will be 450 mm. The resulting wafers are mechanically lapped and ground on both sides to achieve a flat surface, and to remove the mechanical damage due to sawing. Such damage would have a detrimental effect on devices. The flatness of the wafer is critical from the point of view of “depth of focus” or how sharp an image can be focussed on the wafer surface during photolithography, as discussed in Chapter 5. The Si wafers are then rounded or “chamfered” along the edges to minimize the likelihood of chipping the wafers during processing. Finally, the wafers undergo chemical–mechanical





**Figure 1–12**

Steps involved in manufacturing Si wafers. (Photograph courtesy of MEMC Electronics Intl.)

polishing using a slurry of very fine  $\text{SiO}_2$  particles in a basic  $\text{NaOH}$  solution to give the front surface of the wafer a mirror-like finish. The wafers are now ready for IC fabrication (Fig. 1–12). The economic value added in this process is impressive. From sand ( $\text{SiO}_2$ ) costing pennies, we can obtain Si wafers costing a few hundred dollars, on which we can make hundreds of microprocessors, for example, each costing several hundred dollars.

### 1.3.4 Doping

As previously mentioned, there are some impurities in the molten EGS. We may also add intentional impurities or dopants to the Si melt to change its electronic properties. At the solidifying interface between the melt and the solid, there will be a certain distribution of impurities between the two phases. An important quantity that identifies this property is the *distribution coefficient*  $k_d$ , which is the ratio of the concentration of the impurity in the solid  $C_S$  to the concentration in the liquid  $C_L$  at equilibrium:

$$k_d = \frac{C_S}{C_L} \quad (1-6)$$

The distribution coefficient is a function of the material, the impurity, the temperature of the solid–liquid interface, and the growth rate. For an impurity with a distribution coefficient of one-half, the relative concentration of the impurity in the molten liquid to that in the refreezing solid is two to one. Thus the concentration of impurities in that portion of material that solidifies first is one-half the original concentration  $C_0$ . The distribution coefficient is thus important during growth from a melt. This can be illustrated by an example involving Czochralski growth:

*Find the weight of As ( $k_d = 0.3$ ) added to 1 kg Si in Czochralski growth for  $10^{15} \text{ cm}^{-3}$  doping.*

#### EXAMPLE 1–4

$$\text{atomic weight of As} = 74.9 \frac{\text{g}}{\text{mol}}$$

#### SOLUTION

$$C_s = k_d \cdot C_L = 10^{15} \frac{1}{\text{cm}^3} \rightarrow C_L \frac{10^{15} \frac{1}{\text{cm}^3}}{0.3} = 3.33 \cdot 10^{15} \frac{1}{\text{cm}^3}$$

Assume As may be neglected for overall melt weight and volume

$$\frac{1000 \text{ g Si}}{2.33 \frac{\text{g}}{\text{cm}^3}} = 429.2 \text{ cm}^3 \text{ Si}$$

$$3.33 \cdot 10^{15} \frac{1}{\text{cm}^3} \cdot 429.2 \text{ cm}^3 = 1.43 \cdot 10^{18} \text{ As atoms}$$

$$\frac{1.43 \cdot 10^{18} \text{ atoms} \cdot 74.9 \frac{\text{g}}{\text{mol}}}{6.02 \cdot 10^{23} \frac{\text{atoms}}{\text{mol}}} = 1.8 \cdot 10^{-4} \text{ g As} = 1.8 \cdot 10^{-7} \text{ kg As}$$

One of the most important and versatile methods of crystal growth for device applications is the growth of a thin crystal layer on a wafer of a compatible crystal. The substrate crystal may be a wafer of the same material as the grown layer or a different material with a similar lattice structure. In this process the substrate serves as the seed crystal onto which the new crystalline material grows. The growing crystal layer maintains the crystal structure and orientation of the substrate. The technique of growing an oriented single-crystal layer on a substrate wafer is called *epitaxial growth*, or *epitaxy*. As we shall see in this section, epitaxial growth can be performed at temperatures considerably below the melting point of the substrate crystal. A variety of methods are used to provide the appropriate atoms to the surface of the growing layer. These methods include *chemical vapor deposition (CVD)*,<sup>2</sup>

#### 1.4 EPITAXIAL GROWTH

<sup>2</sup>The generic term *chemical vapor deposition* includes the deposition of layers that may be polycrystalline or amorphous. When a CVD process results in a single-crystal epitaxial layer, a more specific term is *vapor-phase epitaxy (VPE)*.

growth from a melt (*liquid-phase epitaxy, LPE*), and evaporation of the elements in a vacuum (*molecular beam epitaxy, MBE*). With this wide range of epitaxial growth techniques, it is possible to grow a variety of crystals for device applications, having properties specifically designed for the electronic or optoelectronic device being made.

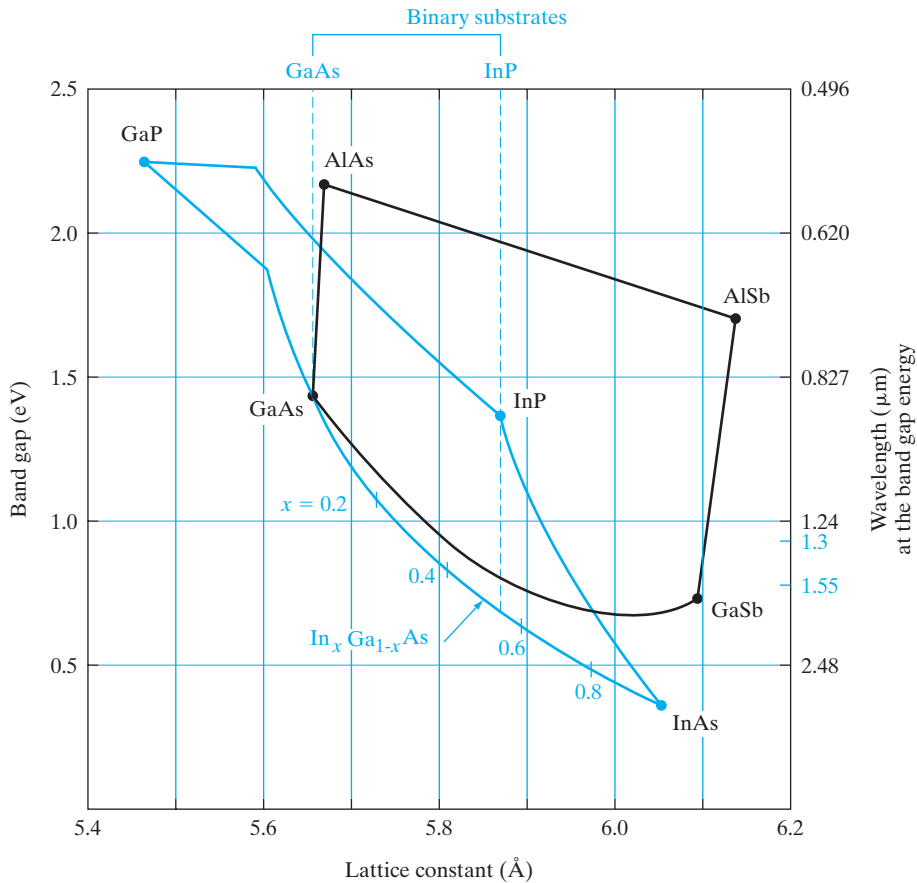
### 1.4.1 Lattice-Matching in Epitaxial Growth

When Si epitaxial layers are grown on Si substrates, there is a natural matching of the crystal lattice, and high-quality single-crystal layers result. On the other hand, it is often desirable to obtain epitaxial layers that differ somewhat from the substrate, which is known as *heteroepitaxy*. This can be accomplished easily if the lattice structure and lattice constant  $a$  match for the two materials. For example, GaAs and AlAs both have the zinc blende structure, with a lattice constant of about 5.65 Å. As a result, epitaxial layers of the ternary alloy AlGaAs can be grown on GaAs substrates with little lattice mismatch. Similarly, GaAs can be grown on Ge substrates (see Appendix III).

Since AlAs and GaAs have similar lattice constants, it is also true that the ternary alloy AlGaAs has essentially the same lattice constant over the entire range of compositions from AlAs to GaAs. As a result, one can choose the composition  $x$  of the ternary compound  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  to fit the particular device requirement, and grow this composition on a GaAs wafer. The resulting epitaxial layer will be lattice-matched to the GaAs substrate.

Figure 1–13 illustrates the energy band gap  $E_g$  as a function of lattice constant  $a$  for several III–V ternary compounds as they are varied over their composition ranges. For example, as the ternary compound InGaAs is varied by choice of composition on the column III sublattice from InAs to GaAs, the band gap changes from 0.36 to 1.43 eV while the lattice constant of the crystal varies from 6.06 Å for InAs to 5.65 Å for GaAs. Clearly, we cannot grow this ternary compound over the entire composition range on a particular binary substrate, which has a fixed lattice constant. As Fig. 1–13 illustrates, however, it is possible to grow a specific composition of InGaAs on an InP substrate. The vertical (invariant lattice constant) line from InP to the InGaAs curve shows that a midrange ternary composition (actually,  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ ) can be grown lattice-matched to an InP substrate. Similarly, a ternary InGaP alloy with about 50 percent Ga and 50 percent In on the column III sublattice can be grown lattice-matched to a GaAs substrate. To achieve a broader range of alloy compositions, grown lattice-matched on particular substrates, it is helpful to use quaternary alloys such as InGaAsP. The variation of compositions on both the column III and column V sublattices provides additional flexibility in choosing a particular band gap while providing lattice-matching to convenient binary substrates such as GaAs or InP.

In the case of GaAsP, the lattice constant is intermediate between that of GaAs and GaP, depending upon the composition. For example, GaAsP crystals used in red LEDs have 40 percent phosphorus and 60 percent arsenic on the column V sublattice. Since such a crystal cannot be grown directly on

**Figure 1-13**

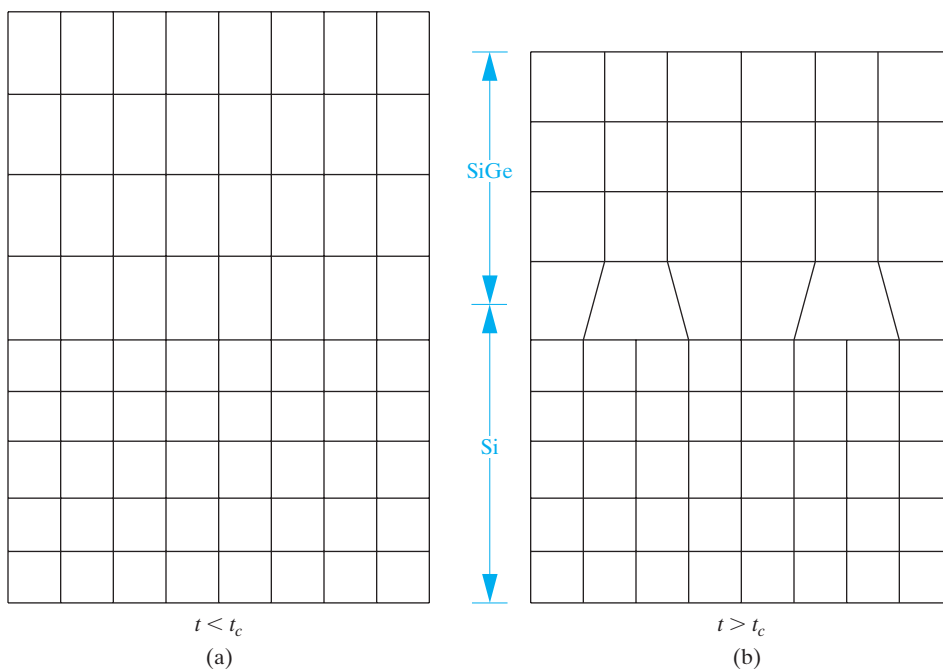
Relationship between band gap and lattice constant for alloys in the InGaAsP and AlGaAsSb systems. The dashed vertical lines show the lattice constants for the commercially available binary substrates GaAs and InP. For the marked example of  $\text{In}_x\text{Ga}_{1-x}\text{As}$ , the ternary composition  $x = 0.53$  can be grown lattice-matched on InP, since the lattice constants are the same. For quaternary alloys, the compositions on both the III and V sublattices can be varied to grow lattice-matched epitaxial layers along the dashed vertical lines between curves. For example,  $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{P}_{1-y}$  can be grown on InP substrates, with resulting band gaps ranging from 0.75 eV to 1.35 eV. In using this figure, assume the lattice constant  $a$  of a ternary alloy varies linearly with the composition  $x$ .

either a GaAs or a GaP substrate, it is necessary to gradually change the lattice constant as the crystal is grown. Using a GaAs or Ge wafer as a substrate, the growth is begun at a composition near GaAs. A region  $\sim 25\ \mu\text{m}$  thick is grown while gradually introducing phosphorus until the desired As/P ratio is achieved. The desired epitaxial layer (e.g.,  $100\ \mu\text{m}$  thick) is then grown on this graded layer. By this method epitaxial growth always occurs on a crystal of similar lattice constant. Although some crystal dislocations occur due to lattice strain in the graded region, such crystals are of high quality and can be used in LEDs.

In addition to the widespread use of lattice-matched epitaxial layers, the advanced epitaxial growth techniques described in the following sections allow the growth of very thin ( $\sim 100$  Å) layers of lattice-mismatched crystals. If the mismatch is only a few percent and the layer is thin, the epitaxial layer grows with a lattice constant in compliance with that of the seed crystal (Fig. 1–14). The resulting layer is in compression or tension along the surface plane as its lattice constant adapts to the seed crystal (Fig. 1–14). Such a layer is called *pseudomorphic* because it is not lattice-matched to the substrate without strain. However, if the epitaxial layer exceeds a critical layer thickness,  $t_c$ , which depends on the lattice mismatch, the strain energy leads to formation of defects called *misfit dislocations*. Using thin alternating layers of slightly mismatched crystal layers, it is possible to grow a *strained-layer superlattice (SLS)* in which alternate layers are in tension and compression. The overall SLS lattice constant is an average of that of the two bulk materials.

### 1.4.2 Vapor-Phase Epitaxy

The advantages of low temperature and high-purity epitaxial growth can be achieved by crystallization from the vapor phase. Crystalline layers can be grown onto a seed or substrate from a chemical vapor of the semiconductor material



**Figure 1–14**

Heteroepitaxy and misfit dislocations. For example, in heteroepitaxy of a SiGe layer on Si, the lattice mismatch between SiGe and Si leads to compressive strain in the SiGe layer. The amount of strain depends on the mole fraction of Ge. (a) For layer thicknesses less than the critical layer thickness,  $t_c$ , pseudomorphic growth occurs. (b) However, above  $t_c$ , misfit dislocations form at the interface which may reduce the usefulness of the layers in device applications.

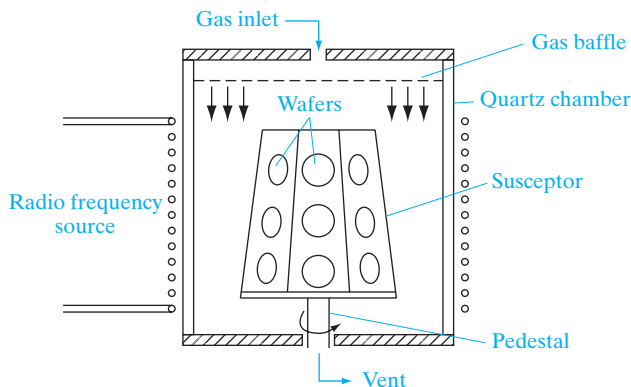
or from mixtures of chemical vapors containing the semiconductor. *Vapor-phase epitaxy (VPE)* is a particularly important source of semiconductor material for use in devices. Some compounds such as GaAs can be grown with better purity and crystal perfection by vapor epitaxy than by other methods. Furthermore, these techniques offer great flexibility in the actual fabrication of devices. When an epitaxial layer is grown on a substrate, it is relatively simple to obtain a sharp demarcation between the type of impurity doping in the substrate and in the grown layer. The advantages of this freedom to vary the impurity will be discussed in subsequent chapters. We point out here, however, that Si IC devices (Chapter 9) are usually built in layers grown by VPE on Si wafers.

Epitaxial layers are generally grown on Si substrates by the controlled deposition of Si atoms onto the surface from a chemical vapor containing Si. In one method, a gas of silicon tetrachloride reacts with hydrogen gas to give Si and anhydrous HCl:



If this reaction occurs at the surface of a heated crystal, the Si atoms released in the reaction can be deposited as an epitaxial layer. The HCl remains gaseous at the reaction temperature and does not disturb the growing crystal. As indicated, this reaction is reversible. This is very important because it implies that by adjusting the process parameters, the reaction in Eq. (1-7) can be driven to the left (providing etching of the Si rather than deposition). This etching can be used for preparing an atomically clean surface on which epitaxy can occur.

This vapor epitaxy technique requires a chamber into which the gases can be introduced and a method for heating the Si wafers. Since the chemical reactions take place in this chamber, it is called a *reaction chamber* or, more simply, a *reactor*. Hydrogen gas is passed through a heated vessel in which  $\text{SiCl}_4$  is evaporated; then the two gases are introduced into the reactor over the substrate crystal, along with other gases containing the desired doping impurities. The Si slice is placed on a graphite susceptor or some other material that can be heated to the reaction temperature with an rf heating coil or tungsten halogen lamps. This method can be adapted to grow epitaxial layers of closely controlled impurity concentration on many Si slices simultaneously (Fig. 1-15).



**Figure 1-15**

A barrel-type reactor for Si vapor-phase epitaxy. These are atmospheric pressure systems. The Si wafers are held in slots cut on the sides of a SiC-coated graphite susceptor that flares out near the base to promote gas flow patterns conducive to uniform epitaxy.

The reaction temperature for the hydrogen reduction of  $\text{SiCl}_4$  is approximately 1150–1250°C. Other reactions may be employed at somewhat lower temperatures, including the use of dichlorosilane ( $\text{SiH}_2\text{Cl}_2$ ) at 1000–1100°C, or the pyrolysis of silane ( $\text{SiH}_4$ ) at 500–1000°C. Pyrolysis involves the breaking up of the silane at the reaction temperature:



There are several advantages of the lower reaction temperature processes, including the fact that they reduce migration of impurities from the substrate to the growing epitaxial layer.

In some applications it is useful to grow thin Si layers on insulating substrates. For example, vapor-phase epitaxial techniques can be used to grow ~1  $\mu\text{m}$  Si films on sapphire and other insulators.

Vapor-phase epitaxial growth is also important in the III–V compounds, such as GaAs, GaP, and the ternary alloy GaAsP, which is widely used in the fabrication of LEDs. Substrates are held at about 800°C on a rotating wafer holder while phosphine, arsine, and gallium chloride gases are mixed and passed over the samples. The GaCl is obtained by reacting anhydrous HCl with molten Ga within the reactor. Variation of the crystal composition for GaAsP can be controlled by altering the mixture of arsine and phosphine gases.

Another useful method for epitaxial growth of compound semiconductors is called *metal-organic vapor-phase epitaxy* (MOVPE), or *organo-metallic vapor-phase epitaxy* (OMVPE). For example, the organometallic compound trimethylgallium can be reacted with arsine to form GaAs and methane:

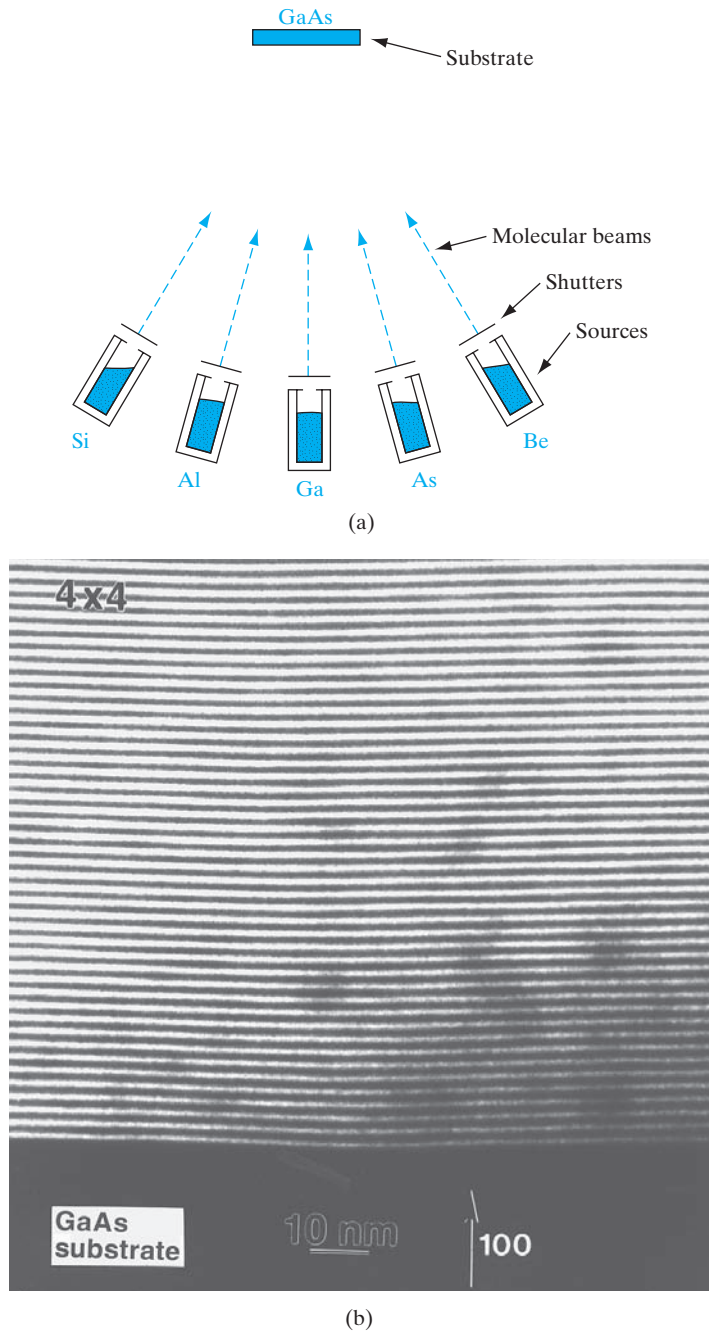


This reaction takes place at about 700°C, and epitaxial growth of high-quality GaAs layers can be obtained. Other compound semiconductors can also be grown by this method. For example, trimethylaluminum can be added to the gas mixture to grow AlGaAs. This growth method is widely used in the fabrication of a variety of devices, including solar cells and lasers. The convenient variability of the gas mixture allows the growth of multiple thin layers similar to those discussed below for molecular beam epitaxy.

### 1.4.3 Molecular Beam Epitaxy

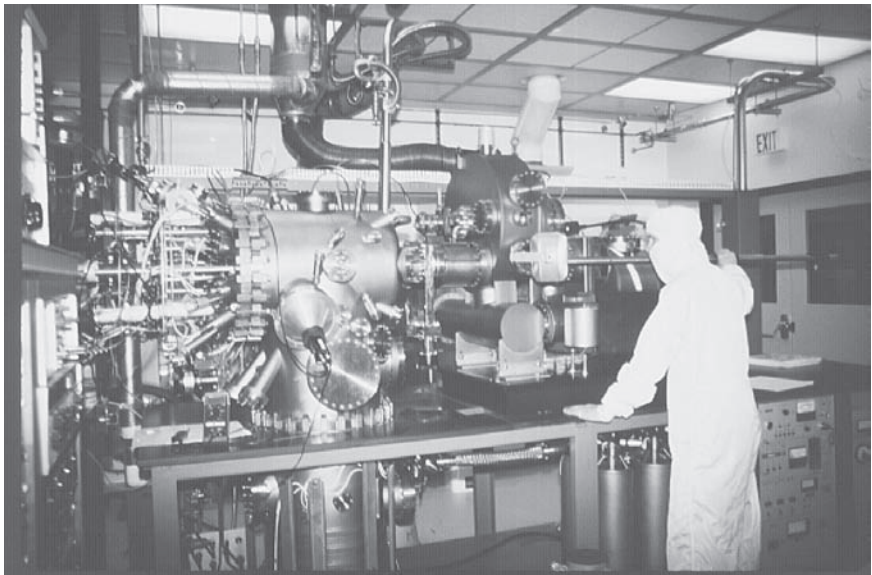
One of the most versatile techniques for growing epitaxial layers is called *molecular beam epitaxy* (MBE). In this method the substrate is held in a high vacuum while molecular or atomic beams of the constituents impinge upon its surface (Fig. 1–16a). For example, in the growth of AlGaAs layers on GaAs substrates, the Al, Ga, and As components, along with dopants, are heated in separate cylindrical cells. Collimated beams of these constituents escape into the vacuum and are directed onto the surface of the substrate.



**Figure 1-16**

Crystal growth by molecular beam epitaxy (MBE): (a) evaporation cells inside a high-vacuum chamber directing beams of Al, Ga, As, and dopants onto a GaAs substrate; (b) scanning electron micrograph of the cross section of an MBE-grown crystal having alternating layers of GaAs (dark lines) and AlGaAs (light lines). Each layer is four monolayers ( $4 \times a/2 = 11.3 \text{ \AA}$ ) thick.

**Figure 1-17**  
Molecular  
beam epitaxy  
facility in the  
Microelectronics  
Research Center  
at the University  
of Texas at Austin.



The rates at which these atomic beams strike the surface can be closely controlled, and growth of very high-quality crystals results. The sample is held at a relatively low temperature (about  $600^{\circ}\text{C}$  for GaAs) in this growth procedure. Abrupt changes in doping or in crystal composition (e.g., changing from GaAs to AlGaAs) can be obtained by controlling shutters in front of the individual beams. Using slow growth rates ( $\leq 1 \mu\text{m/h}$ ), it is possible to control the shutters to make composition changes on the scale of the lattice constant. For example, Fig. 1-16b illustrates a portion of a crystal grown with alternating layers of GaAs and AlGaAs only four monolayers thick. Because of the high vacuum and close controls involved, MBE requires a rather sophisticated setup (Fig. 1-17). However, the versatility of this growth method makes it very attractive for many applications.

As MBE has developed in recent years, it has become common to replace some of the solid sources shown in Fig. 1-16 with gaseous chemical sources. This approach, called *chemical beam epitaxy*, or *gas-source MBE*, combines many of the advantages of MBE and VPE.

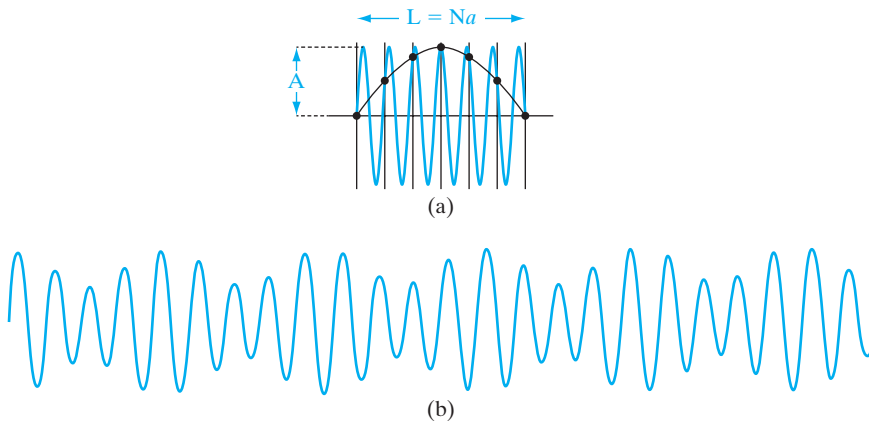
### 1.5 WAVE PROPAGATION IN DISCRETE, PERIODIC STRUCTURES

In the next two chapters, we will discuss the propagation of various types of waves in finite size crystals. It is instructive to look at some of the general attributes of such waves in a medium made of discrete, periodic atoms. Recall that a wave is characterized by various parameters such as wavelength,  $\lambda$ , angular frequency,  $\omega = 2\pi\nu$ , phase,  $\phi$ , and intensity,  $I = A^2$ , where  $A$  is the amplitude of the wave [Fig. 1-18a]. The speed of the wave,  $v_p = \lambda\nu$ . Instead of wavelength, one can also describe a wave in terms of the wavevector,  $k = 2\pi/\lambda$ . The velocity  $v_p$ , which is called the *phase velocity* of the wave, then can also be written as  $\omega/k$ . In general, one may deal with not a single wave train, but a wave packet,

made up of a superposition of various waves [Fig. 1–18b]. The wave packet can be shown to travel with a *group velocity*,  $v_g = d\omega/dk$ . A plane wave,  $\psi$ , propagating in the  $\pm x$  direction can be written as  $\psi = A \exp \{j(kx \pm \omega t)\}$ , where  $j$  is the imaginary number. Equivalently, one can take linear combinations of these plane waves and describe them as standing waves  $\sin(kx \pm \omega t)$  and/or  $\cos(kx \pm \omega t)$ .

There are a few quirks when one looks at plane waves propagating in a finite crystal with lattice constant  $a$  and of length  $L = Na$ , where  $N$  is the number of atoms [Fig. 1–18a]. Plane traveling waves cannot propagate in a finite crystal—reflections at the boundaries will give rise to standing wave patterns. It turns out that the mathematics is simpler in terms of plane waves propagating in an infinite crystal. As long as we are interested in the bulk, rather than the surface, properties of the crystal, one can artificially create an infinite crystal by repeating the finite crystal of length  $L$  over all space, and requiring that the value of the wave is equal on the two faces of the finite crystal,  $\psi(0) = \psi(L)$ . This is known as the *periodic boundary condition* [Fig. 1–18b].

The second quirk about these waves is that in a *discrete* medium such as a crystal, where atoms are located at a spacing  $a$ , the wavelengths are physically distinguishable only up to a point. This is clear from Fig. 1–18a, where one sees that two waves representing the displacement of the atoms are physically indistinguishable. This means that the wavelengths below  $2a$ , or wave vectors,  $k$ , greater than  $2\pi/2a = \pi/a$  are meaningless. The maximum range of  $k$  is known as the Brillouin zone. Notice that this would not be true in a *continuous* medium, that is, where the lattice constant  $a$  approaches zero.



**Figure 1–18**

(a) The displacement pattern of the atoms shown in black in a one-dimensional crystal can be described equivalently by a longer or shorter wavelength. The shortest possible wavelength (i.e., largest  $k$  vector) determines the Brillouin zone. In a finite crystal, we have a standing wave pattern. (b) A more complicated wave involving a mixing of two wavelengths (or Fourier components in  $k$ -space) giving rise to a beat pattern. By using a periodic boundary condition for the finite crystal in (a), one ends up with propagating waves rather than standing waves, making the math simpler.

Problems of wave propagation in periodic structures can be described in so-called *real space*,  $x$ , or alternatively in reciprocal  $k$ -space. An analogy would be describing a periodic signal  $f(t) = f(t + T)$ , either in the time domain or by taking the Fourier transform, in the frequency domain,  $F(\omega)$ . Here, the time period  $T$  is equivalent to the length of the crystal,  $L$ . One can extend this analogy further by considering not continuous signals  $f(t)$ , but sampled time-domain signals as in digital signal processing. If the sampling rate is  $\Delta t$ , one must use a Discrete Fourier Transform, where the highest relevant frequency is the Nyquist frequency. The sampling rate  $\Delta t$  corresponds to the lattice constant,  $a$ , and the Nyquist frequency corresponds to the edge of the Brillouin zone.

Incidentally, in Chapter 3 we will see that such waves describing physical displacement of atoms correspond to sound waves propagating in the crystal, and are called phonons. In Chapters 2 and 3, we will see that electrons propagating in crystals can also be described by a wave using the language of quantum mechanics, in term of a complex “wavefunction.” However, such electron wavefunctions are described by the plane waves mentioned above multiplied by another function  $U(k_x, x) = U(k_x, x+a)$  having the same periodicity as the lattice. Such wavefunctions are known as Bloch functions.

---

## SUMMARY

- 1.1** Semiconductor devices are at the heart of information technology. Elemental semiconductors such as Si appear in column IV of the periodic table, while compound semiconductors such as GaAs consist of elements symmetrically around column IV. More complicated alloy semiconductors are used to optimize optoelectronic properties.
- 1.2** These devices are generally made in single-crystal material for best performance. Single crystals have long-range order, while polycrystalline and amorphous materials have short-range and no order, respectively.
- 1.3** Lattices are determined by symmetry. In 3-D, these are called Bravais lattices. When we put a basis of atom(s) on the lattice sites, we get a crystal. Common semiconductors have an fcc symmetry with a basis of two identical or different atoms, resulting in diamond or zinc blende crystals, respectively.
- 1.4** The fundamental building block of a lattice is a *primitive* cell with lattice points at its corners. Sometimes it is easier to describe the crystal in terms of a larger “unit” cell with lattice points not only at the corners but also at body or face centers.
- 1.5** Translating unit cells by integer numbers of basis vectors can replicate the lattice. Planes and directions in a lattice can be defined in terms of Miller indices.
- 1.6** Real crystals can have defects in 0-, 1-, 2-, and 3-D, some of which are benign, but many of which are harmful for device operation.
- 1.7** Semiconductor bulk crystals are grown from a melt by the Czochralski method, starting from a seed. Single-crystal epitaxial layers can be grown on top of semiconductor wafers in various ways, such as VPE, metal-organic chemical vapor deposition (MOCVD), or MBE. One can thereby optimize doping and band-structure properties for device fabrication.

## PROBLEMS

- 1.1** Using Appendix III, which of the listed semiconductors in Table 1–1 has the largest band gap? The smallest? What are the corresponding wavelengths if light is emitted at the energy  $E_g$ ? Is there a noticeable pattern in the band gap energy of III–V compounds related to the column III element?
- 1.2** For a bcc lattice structure with a lattice constant of  $3 \text{ \AA}$ , calculate the separating distance between the nearest atoms, the radius, and the volume of each atom. Also find the maximum packing fraction.
- 1.3** Label the planes illustrated in Fig. P1–3.

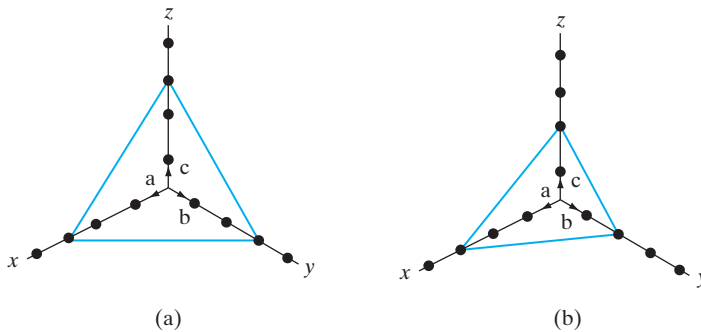


Figure P1–3

- 1.4** Sketch a bcc unit cell with a monoatomic basis. If the atomic density is  $1.6 \times 10^{22} \text{ cm}^{-3}$ , calculate the lattice constant. What is the atomic density per unit area on the (110) plane? What is the radius of each atom? What are interstitials and vacancies?
- 1.5** Calculate the densities of Si and GaAs from the lattice constants (Appendix III), atomic weights, and Avogadro's number. Compare the results with densities given in Appendix III. The atomic weights of Si, Ga, and As are 28.1, 69.7, and 74.9, respectively.
- 1.6** The atomic radius of Ga and As are 136 pm and 114 pm, respectively. Using hard sphere approximation, find the lattice constant of GaAs and the volume of the primitive cell. [ $1 \text{ \AA} = 100 \text{ pm}$ ].
- 1.7** Sketch an fcc lattice unit cell (lattice constant =  $5 \text{ \AA}$ ) with a monoatomic basis, and calculate the atomic density per unit area on (110) planes. What is the atomic density per unit volume? Indicate an interstitial defect in this cell.
- 1.8** Sketch a view down a  $\langle 110 \rangle$  direction of a diamond lattice, using Fig. 1–9 as a guide. Include lines connecting nearest neighbors.
- 1.9** Show by a sketch that the bcc lattice can be represented by two interpenetrating sc lattices. To simplify the sketch, show a  $\langle 100 \rangle$  view of the lattice.
- 1.10** (a) Find the number of atoms/cm<sup>2</sup> on the (100) surface of a Si wafer.  
(b) What is the distance (in  $\text{\AA}$ ) between nearest In neighbors in InP.

- 1.11** The ionic radii of  $\text{Na}^+$  (atomic weight 23) and  $\text{Cl}^-$  (atomic weight 35.5) are 1.0 Å and 1.8 Å, respectively. Treating the ions as hard spheres, calculate the density of NaCl. Compare this with the measured density of 2.17 g/cm<sup>3</sup>.
- 1.12** Sketch an sc unit cell with lattice constant  $a = 4$  Å, whose diatomic basis of atom A is located at the lattice sites, and with atom B displaced by  $(a/2, 0, 0)$ . Assume that both atoms have the same size and we have a close-packed structure (i.e., nearest neighbor atoms touch each other). Calculate
- (i) the packing fraction (i.e., fraction of the total volume occupied by atoms),
  - (ii) the number of B atoms per unit *volume*,
  - (iii) the number of A atoms per unit *area* on (100) planes.
- 1.13** How many atoms are found inside a unit cell of an sc, a bcc, and an fcc crystal? How far apart in terms of lattice constant  $a$  are the nearest neighbor atoms in each case, measured from center to center?
- 1.14** Draw a cube such as Fig. 1–7, and show four {111} planes with different orientations. Repeat for {110} planes.
- 1.15** Find the maximum fractions of the unit cell volume that can be filled by hard spheres in the sc, fcc, and diamond lattices.
- 1.16** Calculate the densities of Ge and InP from the lattice constants (Appendix III), atomic weights, and Avogadro's number. Compare the results with the densities given in Appendix III.
- 1.17** Beginning with a sketch of an fcc lattice, add atoms at  $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$  from each fcc atom to obtain the diamond lattice. Show that only the four added atoms in Fig. 1–8a appear in the diamond unit cell.
- 1.18** Assuming that the lattice constant varies linearly with composition  $x$  for a ternary alloy (e.g., see the variation for InGaAs in Fig. 1–13), what composition of  $\text{AlSb}_x\text{As}_{1-x}$  is lattice-matched to InP? What composition of  $\text{In}_x\text{Ga}_{1-x}\text{P}$  is lattice-matched to GaAs? What is the band gap energy in each case?
- [Note: Such linear variations of crystal properties (e.g., lattice constant and band gap) with mole fraction in alloys is known as *Vegard's law*. A second-order polynomial or quadratic fit to the data is called the *bowing* parameter.]
- 1.19** (a) Find the composition of  $\text{In}_{1-x}\text{Ga}_x\text{As}$  grown lattice-matched on InP substrate. The lattice constants are:  $a(\text{InAs}) = 6.0584$  Å,  $a(\text{GaAs}) = 5.6533$  Å, and  $a(\text{InP}) = 5.8688$  Å.
- (b) An alloy of  $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$  is grown pseudomorphically on a GaAs substrate. Determine the maximum thickness of the grown layer.

---

**READING LIST** **Ashcroft, N. W., and N. D. Mermin.** *Solid State Physics*. Philadelphia: W.B. Saunders, 1976.

**Kittel, C.** *Introduction to Solid State Physics*, 7th ed. New York: Wiley, 1996.

**Plummer, J. D., M. D. Deal, and P. B. Griffin.** *Silicon VLSI Technology*. Upper Saddle River, NJ: Prentice Hall, 2000.

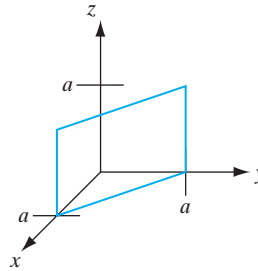
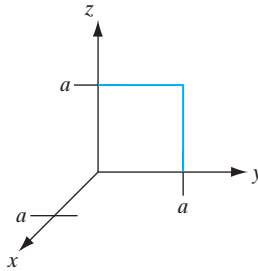
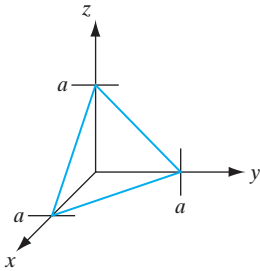
**Stringfellow, G. B.** *Organometallic Vapor-Phase Epitaxy*. New York: Academic Press, 1989.

**Swaminathan, V., and A. T. Macrander.** *Material Aspects of GaAs and InP Based Structures*. Englewood Cliffs, NJ: Prentice Hall, 1991.

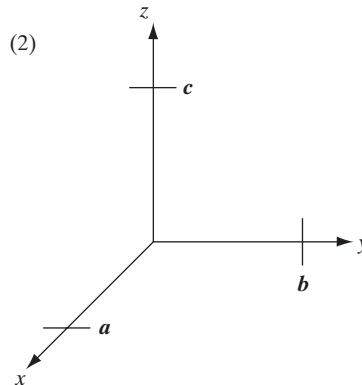
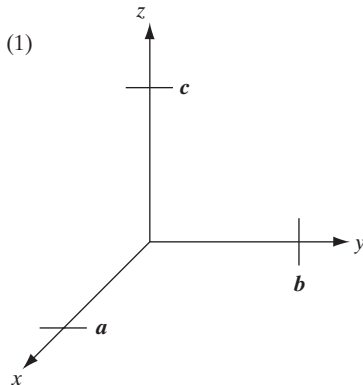
### Question 1

### SELF QUIZ

- (a) Label the following planes using the correct notation for a cubic lattice of unit cell edge length  $a$  (shown within the unit cell).



- (b) Write out all of the equivalent  $\langle 100 \rangle$  directions using the correct notation.  
 (c) On the two following sets of axes, (1) sketch the  $[011]$  direction and (2) a  $(111)$  plane (for a cubic system with primitive vectors  $\mathbf{a}$ ,  $\mathbf{b}$ , and  $\mathbf{c}$ ).

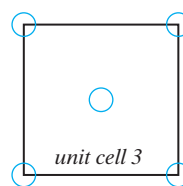
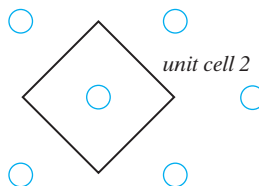
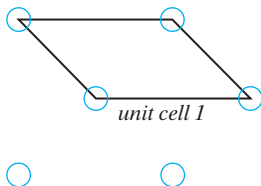


### Question 2

- (a) Which of the following three unit cells are *primitive cells* for the two-dimensional lattice? Circle the correct combination **in bold** below.

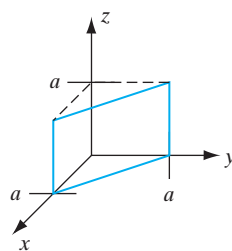
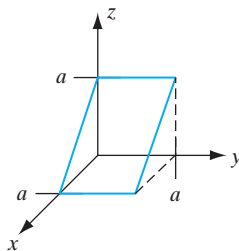
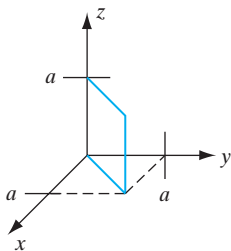
**1 / 2 / 3 / 1 and 2 / 1 and 3 / 2 and 3 / 1, 2, and 3**

**answer:**

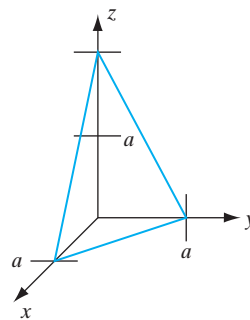
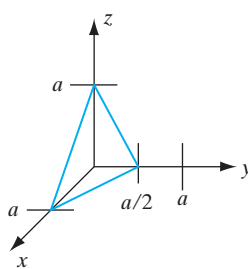
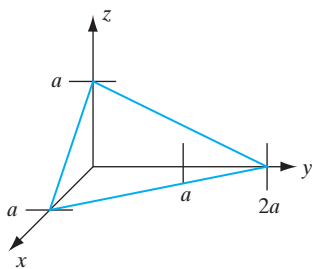




- (b) The following planes (shown within the first quadrant for  $0 < x, y, z < a$  only, with the dotted lines for reference only) are all from what one set of *equivalent* planes? Use correct notation:

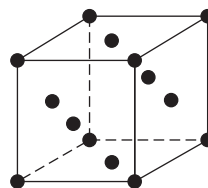
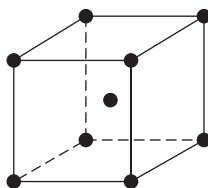
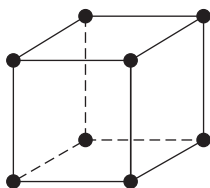


- (c) Which of the following three planes (shown within the first quadrant only) is a (121) plane? Circle the correct diagram.



### Question 3

- (a) Diamond and zinc blende crystal structures are both composed of a Bravais lattice with a two-atom basis. Circle the correct unit cell for this Bravais lattice.



- (b) Which statement below is true?
1. GaAs has a ***diamond / zinc blende*** crystal structure.
  2. Si has a ***diamond / zinc blende*** crystal structure.

**Question 4**

Give some examples of zero-dimensional, one-dimensional, two-dimensional, and three-dimensional defects in a semiconductor.

**Question 5**

- (a) What is the difference between a primitive cell and a unit cell? What is the utility of both concepts?
- (b) What is the difference between a lattice and a crystal? How many different one-dimensional lattices can you have?

**Question 6**

Consider growing InAs on the following crystal substrates: InP, AlAs, GaAs, and GaP. For which case would the *critical thickness* of the InAs layer be greatest? You may use Fig. 1–13 from your text.

***GaP / GaAs / AlAs / InP***

---

## Chapter 2

# Atoms and Electrons

---

### OBJECTIVES

1. Understand the wave-particle duality of nature in quantum mechanics
2. Study the Bohr model of atoms
3. Apply the Schrödinger equation to simple problems
4. Understand the electronic structure of atoms and the periodic table
5. Understand how semiconductor properties are determined

Since this book is primarily an introduction to solid state devices, it would be preferable not to delay this discussion with subjects such as atomic theory, quantum mechanics, and electron models. However, the behavior of solid state devices is directly related to these subjects. For example, it would be difficult to understand how an electron is transported through a semiconductor device without some knowledge of the electron and its interaction with the crystal lattice. Therefore, in this chapter we shall investigate some of the important properties of electrons, with special emphasis on two points: (1) the electronic structure of atoms, and (2) the interaction of atoms and electrons with excitation, such as the absorption and emission of light. By studying electron energies in an atom, we lay the foundation for understanding the influence of the lattice on electrons participating in current flow through a solid. Our discussions concerning the interaction of light with electrons form the basis for later descriptions of changes in the conductivity of a semiconductor with optical excitation, properties of light-sensitive devices, and lasers.

First, we shall investigate some of the experimental observations which led to the modern concept of the atom, and then we shall give a brief introduction to the theory of quantum mechanics. Several important concepts will emerge from this introduction: the electrons in atoms are restricted to certain energy levels by quantum rules; the electronic structure of atoms is determined from these quantum conditions; and this “quantization” defines certain allowable transitions involving absorption and emission of energy by the electrons.

## 2.1 INTRODUCTION TO PHYSICAL MODELS

The main effort of science is to describe what happens in nature, in as complete and concise a form as possible. In physics this effort involves observing natural phenomena, relating these observations to previously established theory, and finally establishing a physical model for the observations. For example, we can explain the behavior of a spring-supported weight moving up and down periodically after an initial displacement, because the differential equations describing such a simple harmonic motion have been established by Newtonian classical mechanics.

When a new physical phenomenon is observed, it is necessary to find out how it fits into the established models and “laws” of physics. In the vast majority of cases this involves a direct extension of the mathematics of well-established models to the particular conditions of the new problem. In fact, it is not uncommon for a scientist or engineer to predict that a new phenomenon should occur before it is actually observed, simply by a careful study and extension of existing models and laws. The beauty of science is that natural phenomena are not isolated events but are related to other events by a few analytically describable laws. However, it does happen occasionally that a set of observations cannot be described in terms of existing theories. In such cases it is necessary to develop models which are based as far as possible on existing laws, but which contain new aspects arising from the new phenomena. Postulating new physical principles is a serious business, and it is done only when there is no possibility of explaining the observations with established theory. When new assumptions and models are made, their justification lies in the following question: “Does the model describe precisely the observations, and can reliable predictions be made based on the model?” The model is good or poor depending on the answer to this question.

In the 1920s it became necessary to develop a new theory to describe phenomena on the atomic scale. A long series of careful observations had been made that clearly indicated that many events involving electrons and atoms did not obey the classical laws of mechanics. It was necessary, therefore, to develop a new kind of mechanics to describe the behavior of particles on this small scale. This new approach, called *quantum mechanics*, describes atomic phenomena very well and also properly predicts the way in which electrons behave in solids—our primary interest here. Through the years, quantum mechanics has been so successful that now it stands beside the classical laws as a valid description of nature.

A special problem arises when students first encounter the theory of quantum mechanics. The problem is that quantum concepts are largely mathematical in nature and do not involve the “common sense” quality associated with classical mechanics. At first, many students find quantum concepts difficult, not so much because of the mathematics involved, but because they feel the concepts are somehow divorced from “reality.” This is a reasonable reaction, since ideas which we consider to be real or intuitively satisfying are usually based on our own observation. Thus the classical laws of motion are easy to understand because we observe bodies in motion every day. On the other hand, we observe the effects of atoms and electrons only indirectly,

and naturally we have very little feeling for what is happening on the atomic scale. It is necessary, therefore, to depend on the facility of the theory to predict experimental results rather than to attempt to force classical analogs onto the nonclassical phenomena of atoms and electrons.

Our approach in this chapter will be to investigate the important experimental observations that led to the quantum theory, and then to indicate how the theory accounts for these observations. Discussions of quantum theory must necessarily be largely qualitative in such a brief presentation, and those topics that are most important to solid state theory will be emphasized here. Several good references for further individual study are given at the end of this chapter.

## 2.2 EXPERIMENTAL OBSERVATIONS

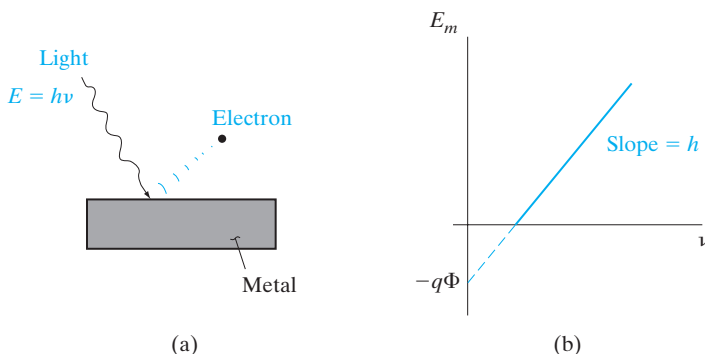
The experiments that led to the development of quantum theory were concerned with the interaction of light and matter. On the one hand, there were phenomena such as interference or diffraction which clearly indicated that light has a wave character as proposed by Huygens, in contrast to the particle or *corpuscular* view of light proposed by Newton. But on the other hand, many experiments at the turn of the 20<sup>th</sup> century clearly showed that a new theory of light was needed.

### 2.2.1 The Photoelectric Effect

An important observation by Planck indicated that radiation from a heated sample, known as blackbody radiation, is emitted in discrete units of energy called *quanta*; the energy units were described by  $h\nu$ , where  $\nu$  is the frequency of the radiation, and  $h$  is a quantity now called Planck's constant ( $h = 6.63 \times 10^{-34}$  J-s). Soon after Planck developed this hypothesis, Einstein interpreted an important experiment that clearly demonstrated the discrete nature (*quantization*) of light. This experiment involved absorption of optical energy by the electrons in a metal and the relationship between the amount of energy absorbed and the frequency of the light (Fig. 2-1). Let us suppose that monochromatic light is incident on the surface of a metal plate in a vacuum. The electrons in the metal absorb energy from the light, and some of the electrons receive enough energy to be ejected from the metal surface into

**Figure 2-1**

The photoelectric effect: (a) electrons are ejected from the surface of a metal when exposed to light of frequency  $\nu$  in a vacuum; (b) plot of the maximum kinetic energy of ejected electrons vs. frequency of the incoming light.



the vacuum. This phenomenon is called the *photoelectric effect*. If the energy of the escaping electrons is measured, a plot can be made of the maximum energy as a function of the frequency  $\nu$  of the incident light (Fig. 2-1b).

One simple way of finding the maximum energy of the ejected electrons is to place another plate above the one shown in Fig. 2-1a and then create an electric field between the two plates. The potential necessary to retard all electron flow between the plates gives the energy  $E_m$ . For a particular frequency of light incident on the sample, a maximum energy  $E_m$  is observed for the emitted electrons. The resulting plot of  $E_m$  vs.  $\nu$  is linear, with a slope equal to Planck's constant. The equation of the line shown in Fig. 2-1b is

$$E_m = h\nu - q\Phi \quad (2-1)$$

where  $q$  is the magnitude of the electronic charge. The quantity  $\Phi$  (volts) is a characteristic of the particular metal used. When  $\Phi$  is multiplied by the electronic charge, an energy (joules) is obtained which represents the minimum energy required for an electron to escape from the metal into a vacuum. The energy  $q\Phi$  is called the *work function* of the metal. These results indicate that the electrons receive an energy  $h\nu$  from the light and lose an amount of energy  $q\Phi$  in escaping from the surface of the metal.

This experiment demonstrates clearly that Planck's hypothesis was correct—light energy is contained in discrete units rather than in a continuous distribution of energies. Other experiments also indicate that, in addition to the wave nature of light, the quantized units of light energy can be considered as localized packets of energy, called *photons*. (Interestingly, this is reminiscent of the Newtonian picture.) We get the Planck relationship

$$E = h\nu = (h/2\pi)(2\pi\nu) = \hbar\omega \quad (2-2a)$$

The photoelectron energy was found *not* to increase with increasing light intensity; instead the *number* of photoelectrons increases. This is in sharp contrast to what classical physics would have predicted. A higher amplitude or intensity wave packs more energy, and should have increased the photoelectron energy. Instead, we find in quantum physics that a higher intensity light corresponds to a larger *number* of these photons, and hence a larger *number* of photoelectrons. Some experiments emphasize the wave nature of light, while other experiments reveal the discrete nature of photons. This wave-particle duality is fundamental to quantum processes and does not imply an ambiguity in the theory.

This duality is seen beautifully in the Young's double slit diffraction experiment. We may recall that if monochromatic light is shone through two narrow slits onto a screen, we see diffraction and interference patterns on the screen corresponding to high intensity bright regions (constructive interference) and low intensity dark regions (destructive interference). This is consistent with the wave picture. Now, if the intensity of the light source is reduced, so that just a few photons are emitted per second, we find that we do not get an interference pattern. Instead, we get specks of light wherever the photon hits the screen, because a photon cannot be split. There is a certain amount of randomness in terms of where the photon hits. However, if we continue

with the experiment, these random hits of the photons gradually build up the interference pattern, corresponding to the wave picture. Statistically, there are more photon hits in the bright regions of the screen, and fewer hits in the dark spots. Similar results were subsequently observed with electrons and other subatomic particles. This wave–particle duality is fundamental to the development of quantum mechanics.

Hence, based on the wave–particle duality of light, Louis de Broglie proposed that particles of matter (such as electrons) similarly could manifest a wave character in certain experiments. This observation was confirmed by the diffraction of electrons by the periodic array of atoms in a crystal observed by Davisson and Germer. De Broglie proposed that a particle of momentum  $p = mv$  has a wavelength given by

$$\begin{aligned}\lambda &= h/p = h/mv \\ \Rightarrow p &= h/\lambda = (h/2\pi)(2\pi/\lambda) = \hbar k\end{aligned}\tag{2-2b}$$

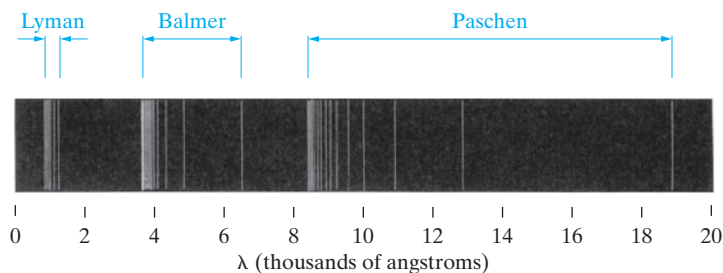
The Planck and de Broglie relationships are fundamental in quantum physics and are valid for *all* situations and objects, including photons and electrons. They connect the wave description of phenomena (frequency and wavelength) to a particle description (energy and momentum).

The relation between frequency and wavelength (or equivalently, energy and momentum), known as the *dispersion relationship*, however, is *not* the same for different objects. For example, for photons, the wavelength ( $\lambda$ ) is related to frequency by  $\nu = c/\lambda \Rightarrow E = \hbar\omega = \hbar(2\pi\nu) = \hbar 2\pi(c/\lambda) = \hbar ck = cp$ , where  $c$  is the speed of light. For electrons, it is an approximately parabolic dispersion relationship,  $E(k)$ , also known as the band structure, as discussed in Chapter 3.

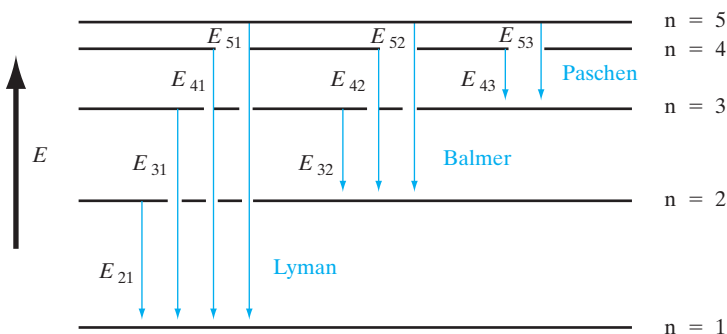
### 2.2.2 Atomic Spectra

One of the most valuable experiments of modern physics is the analysis of absorption and emission of light by atoms. The de Broglie wave property of electrons is key to understanding these experiments. For example, an electric discharge can be created in a gas, so that the atoms begin to emit light with wavelengths characteristic of the gas. We see this effect in a neon sign, which is typically a glass tube filled with neon or a gas mixture, with electrodes for creating a discharge. If the intensity of the emitted light is measured as a function of wavelength, one finds a series of sharp lines rather than a continuous distribution of wavelengths. By the early 1900s the characteristic spectra for several atoms were well known. A portion of the measured emission spectrum for hydrogen is shown in Fig. 2-2, in which the vertical lines represent the positions of observed emission peaks on the wavelength scale. Photon energy  $h\nu$  is then related to wavelength by the dispersion relation  $\lambda = c/\nu$ .

The lines in Fig. 2-2 appear in several groups labeled the *Lyman*, *Balmer*, and *Paschen* series after their early investigators. Once the hydrogen spectrum was established, scientists noticed several interesting relationships among the lines. The various series in the spectrum were observed to follow certain empirical forms:

**Figure 2-2**

Some important lines in the emission spectrum of hydrogen.

**Figure 2-3**

Relationships among photon energies in the hydrogen spectrum.

$$\text{Lyman: } \nu = cR \left( \frac{1}{1^2} - \frac{1}{n^2} \right), \mathbf{n} = 2, 3, 4, \dots \quad (2-3a)$$

$$\text{Balmer: } \nu = cR \left( \frac{1}{2^2} - \frac{1}{n^2} \right), \mathbf{n} = 3, 4, 5, \dots \quad (2-3b)$$

$$\text{Paschen: } \nu = cR \left( \frac{1}{3^2} - \frac{1}{n^2} \right), \mathbf{n} = 4, 5, 6, \dots \quad (2-3c)$$

where  $R$  is a constant called the Rydberg constant ( $R = 109,678 \text{ cm}^{-1}$ ). If the photon energies  $h\nu$  are plotted for successive values of the integer  $\mathbf{n}$ , we notice that each energy can be obtained by taking sums and differences of other photon energies in the spectrum (Fig. 2-3). For example,  $E_{42}$  in the Balmer series is the difference between  $E_{41}$  and  $E_{21}$  in the Lyman series. This relationship among the various series is called the *Ritz combination principle*. Naturally, these empirical observations stirred a great deal of interest in constructing a comprehensive theory for the origin of the photons given off by atoms.

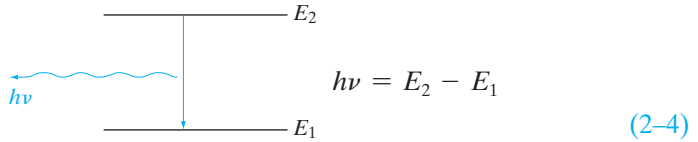
The results of emission spectra experiments led Niels Bohr to construct a model for the hydrogen atom, based on the mathematics of planetary systems. If the electron in the hydrogen atom has a series of planetary-type orbits available to it, it can be excited to an outer orbit and then can fall to any one of the inner orbits, giving off energy corresponding to one of the lines of Fig. 2-3. To develop the model, Bohr made several postulates:

## 2.3

### THE BOHR MODEL



1. Electrons exist in certain stable, circular orbits about the nucleus. This assumption implies that the orbiting electron does not give off radiation as classical electromagnetic theory would normally require of a charge experiencing angular acceleration; otherwise, the electron would not be stable in the orbit but would spiral into the nucleus as it lost energy by radiation.
2. The electron may shift to an orbit of higher or lower energy, thereby gaining or losing energy equal to the difference in the energy levels (by absorption or emission of a photon of energy  $h\nu$ ).

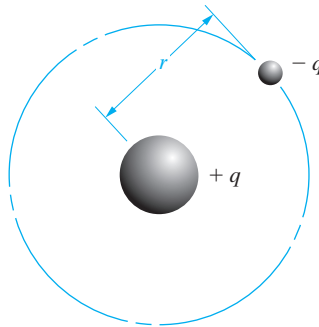


3. The angular momentum  $p_\theta$  of the electron in an orbit is always an integral multiple of Planck's constant divided by  $2\pi$  ( $h/2\pi$  is often abbreviated  $\hbar$  for convenience). This assumption,

$$p_\theta = n\hbar, \quad n = 1, 2, 3, 4, \dots \quad (2-5)$$

is necessary to obtain the observed results of Fig. 2-3. Although Bohr proposed this *ad hoc* relationship simply to explain the data, one can see that this is equivalent to having an integer number of de Broglie wavelengths fit within the circumference of the electron orbit. These were called *pilot waves*, guiding the motion of the electrons around the nucleus. The de Broglie wave concept provided the inspiration for the Schrödinger wave equation in quantum mechanics discussed in Section 2.4.

If we visualize the electron in a stable orbit of radius  $r$  about the proton of the hydrogen atom, we can equate the electrostatic force between the charges to the centripetal force:



$$-\frac{q^2}{Kr^2} = -\frac{mv^2}{r} \quad (2-6)$$

where  $K = 4\pi\epsilon_0$  in MKS units,  $m$  is the mass of the electron, and  $v$  is its velocity. From assumption 3 we have

$$p_\theta = mvr = n\hbar \quad (2-7)$$

Since  $n$  takes on integral values,  $r$  should be denoted by  $r_n$  to indicate the  $n$ th orbit. Then Eq. (2-7) can be written

$$m^2v^2 = \frac{n^2\hbar^2}{r_n^2} \quad (2-8)$$

Substituting Eq. (2-8) in Eq. (2-6) we find that

$$\frac{q^2}{Kr_n^2} = \frac{1}{mr_n} \cdot \frac{n^2\hbar^2}{r_n^2} \quad (2-9)$$

$$r_n = \frac{Kn^2\hbar^2}{mq^2} \quad (2-10)$$

for the radius of the  $n$ th orbit of the electron. Now we must find the expression for the total energy of the electron in this orbit, so that we can calculate the energies involved in transitions between orbits.

From Eqs. (2-7) and (2-10) we have

$$v = \frac{n\hbar}{mr_n} \quad (2-11)$$

$$v = \frac{n\hbar q^2}{Kn^2\hbar^2} = \frac{q^2}{Kn\hbar} \quad (2-12)$$

Therefore, the kinetic energy of the electron is

$$\text{K. E.} = \frac{1}{2}mv^2 = \frac{mq^4}{2K^2n^2\hbar^2} \quad (2-13)$$

The potential energy is the product of the electrostatic force and the distance between the charges:

$$\text{P. E.} = -\frac{q^2}{Kr_n} = -\frac{mq^4}{K^2n^2\hbar^2} \quad (2-14)$$

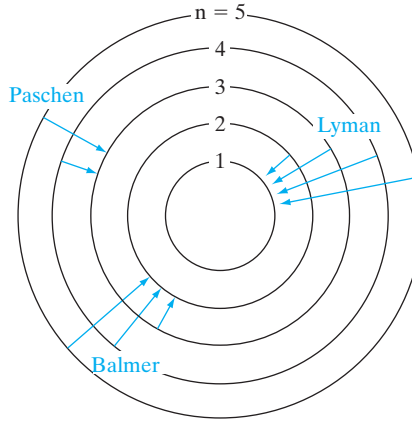
Thus the total energy of the electron in the  $n$ th orbit is

$$E_n = \text{K. E.} + \text{P. E.} = -\frac{mq^4}{2K^2n^2\hbar^2} \quad (2-15)$$

The critical test of the model is whether energy differences between orbits correspond to the observed photon energies of the hydrogen

**Figure 2-4**

Electron orbits and transitions in the Bohr model of the hydrogen atom. Orbit spacing is not drawn to scale.



spectrum. The transitions between orbits corresponding to the Lyman, Balmer, and Paschen series are illustrated in Fig. 2-4. The energy difference between orbits  $n_1$  and  $n_2$  is given by

$$E_{n_2} - E_{n_1} = \frac{mq^4}{2K^2\hbar^2} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad (2-16)$$

The frequency of light given off by a transition between these orbits is

$$\nu_{21} = \left[ \frac{mq^4}{2K^2\hbar^2 h} \right] \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad (2-17)$$

**EXAMPLE 2-1**

Show Eq. (2-17) corresponds to Eq. (2-3). That is, show:  $c \cdot R = \frac{m \cdot q^4}{2 \cdot K^2 \cdot \hbar^2 \cdot h}$ .

**SOLUTION**

From Eq. (2-17) and the solution to Eq. (2-3),

$$\nu_{21} = \frac{c}{\lambda} = \frac{2.998 \cdot 10^8 \frac{\text{m}}{\text{s}}}{9.11 \cdot 10^{-8} \text{ m} \cdot \frac{n_1^2 n_2^2}{n_2^2 - n_1^2}} = 3.29 \cdot 10^{15} \text{ Hz} \cdot \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right)$$

From Eq. (2-3),

$$\begin{aligned} \nu_{21} &= c \cdot R \cdot \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) = 2.998 \cdot 10^8 \frac{\text{m}}{\text{s}} \cdot 1.097 \cdot 10^7 \frac{1}{\text{m}} \cdot \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \\ &= 3.29 \cdot 10^{15} \text{ Hz} \cdot \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \end{aligned}$$

The pre-factor is essentially the Rydberg constant  $R$  times the speed of light  $c$ . A comparison of Eq. (2-17) with the experimental results summed up by Eq. (2-3) indicates that the Bohr theory provides a good model for electronic transitions within the hydrogen atom, as far as the early experimental evidence is concerned.

Whereas the Bohr model accurately describes the gross features of the hydrogen spectrum, it does not include many fine points. For example, experimental evidence indicates some splitting of levels in addition to the levels predicted by the theory. Also, difficulties arise in extending the model to atoms more complicated than hydrogen. Attempts were made to modify the Bohr model for more general cases, but it soon became obvious that a more comprehensive theory was needed. However, the partial success of the Bohr model was an important step toward the eventual development of the quantum theory. The concept that electrons are quantized in certain allowed energy levels, and the relationship of photon energy and transitions between levels, had been established firmly by the Bohr theory.

The principles of quantum mechanics were developed from two different points of view at about the same time (the late 1920s). One approach, developed by Heisenberg, utilizes the mathematics of matrices and is called *matrix mechanics*. Independently, Schrödinger developed an approach utilizing a wave equation, now called *wave mechanics*. These two mathematical formulations appear to be quite different. However, closer examination reveals that beyond the formalism, the basic principles of the two approaches are the same. It is possible to show, for example, that the results of matrix mechanics reduce to those of wave mechanics after mathematical manipulation. We shall concentrate here on the wave mechanics approach, since solutions to a few simple problems can be obtained with it, involving less mathematical discussion.

## 2.4 QUANTUM MECHANICS

### 2.4.1 Probability and the Uncertainty Principle

It is impossible to describe with absolute precision events involving individual particles on the atomic scale. Instead, we must speak of the average values (*expectation values*) of position, momentum, and energy of a particle such as an electron. It is important to note, however, that the uncertainties revealed in quantum calculations are not based on some shortcoming of the theory. In fact, a major strength of the theory is that it describes the probabilistic nature of events involving atoms and electrons. The fact is that such quantities as the position and momentum of an electron *do not exist* apart from a particular uncertainty. The magnitude of this inherent uncertainty is described by the *Heisenberg uncertainty principle*:<sup>1</sup>

<sup>1</sup>This is often called the *principle of indeterminacy*. This is a better term, since the parameters cannot be determined with better accuracy than specified in these equations.

In any measurement of the position and momentum of a particle, the uncertainties in the two measured quantities will be related by

$$(\Delta x)(\Delta p_x) \geq \hbar/2 \quad (2-18)$$

Similarly, the uncertainties in an energy measurement will be related to the uncertainty in the time at which the measurement was made by

$$(\Delta E)(\Delta t) \geq \hbar/2 \quad (2-19)$$

These limitations indicate that simultaneous measurement of position and momentum or of energy and time is inherently inaccurate to some degree. Of course, Planck's constant  $h$  is a rather small number ( $6.63 \times 10^{-34}$  J-s), and we are not concerned with this inaccuracy in the measurement of  $x$  and  $p_x$  for a truck, for example. On the other hand, measurements of the position of an electron and its speed are seriously limited by the uncertainty principle.

One implication of the uncertainty principle is that we cannot properly speak of *the* position of an electron, for example, but must look for the “probability” of finding an electron at a certain position. Thus one of the important results of quantum mechanics is that a *probability density function* can be obtained for a particle in a certain environment, and this function can be used to find the expectation value of important quantities such as position, momentum, and energy. We are familiar with the methods for calculating discrete (single-valued) probabilities from common experience. For example, it is clear that the probability of drawing a particular card out of a random deck is  $1/52$ , and the probability that a tossed coin will come up heads is  $1/2$ . The techniques for making predictions when the probability varies are less familiar, however. In such cases it is common to define a probability of finding a particle within a certain volume. Given a probability density function  $P(x)$  for a one-dimensional problem, the probability of finding the particle in a range from  $x$  to  $x + dx$  is  $P(x)dx$ . Since the particle will be *somewhere*, this definition implies that

$$\int_{-\infty}^{\infty} P(x)dx = 1 \quad (2-20)$$

if the function  $P(x)$  is properly chosen. Equation (2-20) is implied by stating that the function  $P(x)$  is *normalized* (i.e., the integral equals unity).

To find the average value of a function of  $x$ , we need only multiply the value of that function in each increment  $dx$  by the probability of finding the particle in that  $dx$  and sum over all  $x$ . Thus the average value of  $f(x)$  is

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)P(x)dx \quad (2-21a)$$

If the probability density function is not normalized, this equation should be written