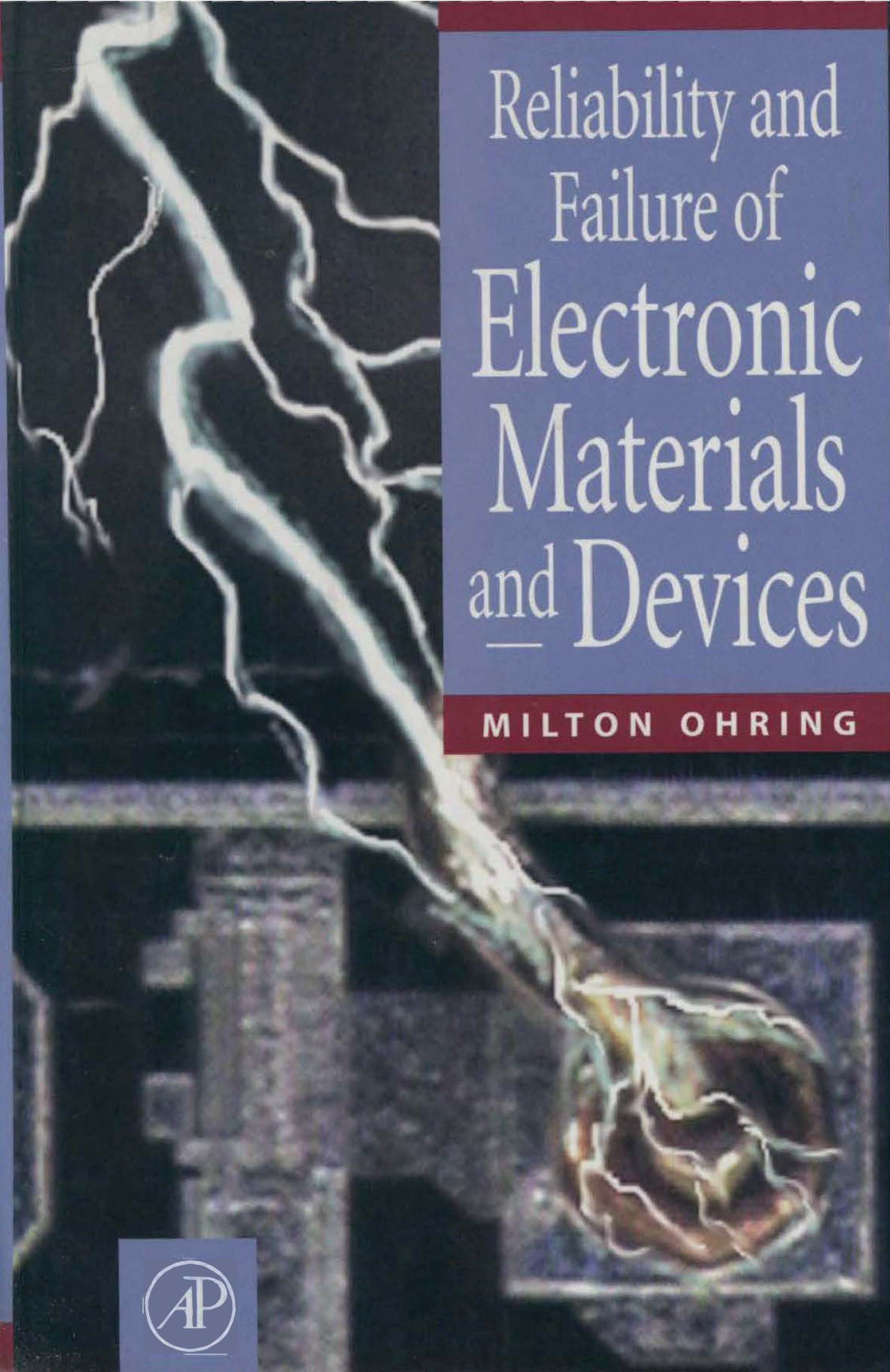




Reliability and Failure of Electronic Materials and Devices

MILTON OHRING



Infineon Exhibit 2001
Innoscence v. Infineon
IPR2024-00975

ENGINEER

89066577594



B89066577594A

Reliability and Failure of Electronic Materials

MILTON OHRING

OHRING

This book introduces the reader to the widely dispersed reliability literature of microelectronic and electro-optical devices. It integrates a treatment of chip and packaging level failures within the contexts of the atomic mechanisms and models used to explain degradation, and the statistical handling of lifetime data. Electromigration, dielectric breakdown, hot-electron effects, electrostatic discharge, corrosion, radiation damage and the mechanical failure of contacts and solder joints are among the failure mechanisms considered. An underlying book thread concerns product defects—their relation to yield and reliability, the role they play in failure, and the way they are experimentally exposed. The book can be used as an advanced undergraduate/graduate textbook for materials scientists and electrical engineers, and as a reference for reliability professionals.

Key Features

- Discusses reliability and failure on both the chip and packaging levels
- Handles the role of defects in yield and reliability
- Includes a tutorial chapter on the mathematics of reliability
- Focuses on electromigration, dielectric breakdown, hot-electron effects, electrostatic discharge, corrosion, radiation damage, and the mechanical failure of packages, contacts, and solder joints
- Considers defect detection methods and failure analysis techniques

About the Author

Dr. Milton Ohring, author of two previously acclaimed Academic Press books, *The Materials Science of Thin Films* (1992) and *Engineering Materials Science* (1995), has taught courses on reliability and failure in electronics at Bell Laboratories (AT&T and Lucent Technologies). From this perspective and the well-written tutorial style of the book, the reader will gain a deeper physical understanding of failure mechanisms in electronic materials and devices; acquire skills in the mathematical handling of reliability data; and better appreciate future technology trends and the reliability issues they raise.

Reliability and Failure of Electronic Materials and Devices



6 08628 24985 8



9 780125 249850

ISBN 0-12-524985-3



90090



TK
7870.23
037
1998

Reliability and Failure of Electronic Materials and Devices

Milton Ohring

*Department of Materials Science and Engineering
Stevens Institute of Technology
Hoboken, New Jersey*



ACADEMIC PRESS

San Diego London Boston
New York Sydney Tokyo Toronto

Kurt K. Wendt Library
University of Wisconsin-Madison
215 N. Randall Avenue
Madison, WI 53706-1688

This book is printed on acid-free paper. ∞

Copyright © 1998 by Academic Press.

All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

The cover image is an artist's rendition of an electrostatic spark discharge event that lasts less than one billionth of a second. This miniature lightning bolt from a charged object can easily damage a semiconductor device. The photo is courtesy of Wayne Tan, Advanced Micro Devices.

ACADEMIC PRESS
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA
1300 Boylston Street, Chestnut Hill, MA 02167, USA
<http://www.apnet.com>

United Kingdom Edition published by
ACADEMIC PRESS LIMITED
24-28 Oval Road, London NW1 7DX
<http://www.hbuk.co.uk/ap>

Library of Congress Cataloging-in-Publication Data

Ohring, Milton, 1936-
Reliability and failure of electronic materials and devices / Milton Ohring.
p. cm.
Includes bibliographical references.
ISBN 0-12-524985-3
1. Electronic apparatus and appliances—Reliability. 2. System failures (Engineering) I. Title.
TK7870.23.037 1998
621.381—dc21

98-16084
C

Printed in the United States of America

98 99 00 01 02 MV 9 8 7 6 5 4 3 2 1

Contents

Acknowledgments	xvii
Preface	xix
<i>Chapter 1</i>	
An Overview of Electronic Devices and Their Reliability	1
1.1 Electronic Products	1
1.1.1 Historical Perspective	1
1.1.2 Solid-State Devices	4
1.1.3 Integrated Circuits	4
1.1.4 Yield of Electronic Products	9
1.2 Reliability, Other “ . . . ilities,” and Definitions	13
1.2.1 Reliability	13
1.2.2 A Brief History of Reliability	14
1.2.3 MIL-HDBK-217	15
1.2.4 Long-Term Nonoperating Reliability	16
1.2.5 Availability, Maintainability, and Survivability	17
1.3 Failure Physics	17
1.3.1 Failure Modes and Mechanisms; Reliable and Failed States	17
1.3.2 Conditions for Change	18
1.3.3 Atom Movements and Driving Forces	20
1.3.4 Failure Times and the Acceleration Factor	24
1.3.5 Load-Strength Interference	25
1.3.6 Semiconductor Device Degradation and Failure	27
1.3.7 Failure Frequency	28
1.3.8 The Bathtub Curve and Failure	29
1.4 Summary and Perspective	31
Exercises	32
References	35

*Chapter 2***Electronic Devices: How They Operate and Are Fabricated . . . 37**

2.1	Introduction	37
2.2	Electronic Materials.	38
2.2.1	Introduction.	38
2.2.2	Semiconductors	39
2.2.3	Conductors	49
2.2.4	Insulators	52
2.3	Diodes	53
2.3.1	The p–n Junction.	53
2.3.2	Contacts	55
2.3.3	Deviations from Ideal Junction Behavior.	57
2.4	Bipolar Transistors	59
2.4.1	Transistors in General	59
2.4.2	Bipolar Junction Transistors	59
2.5	Field Effect Transistors	62
2.5.1	Introduction.	62
2.5.2	The MOS Capacitor.	63
2.5.3	MOS Field Effect Transistor.	66
2.5.4	CMOS Devices	69
2.5.5	MOSFET Instabilities and Malfunction.	70
2.5.6	Bipolar versus CMOS	71
2.5.7	Junction Field Effect Transistor (JFET).	72
2.6	Memories	74
2.6.1	Types of Memories	74
2.6.2	Memories Are Made of This!	76
2.6.3	Reliability Problems in Memories.	77
2.7	GaAs Devices	79
2.7.1	Why Compound Semiconductor Devices?	79
2.7.2	Microwave Applications.	80
2.7.3	The GaAs MESFET.	81
2.7.4	High Electron Mobility Transistor (HEMT).	82
2.7.5	GaAs Integrated Circuits	82
2.8	Electro-optical Devices	84
2.8.1	Introduction.	84
2.8.2	Solar Cells	84
2.8.3	PIN and Avalanche Photodiodes	85
2.8.4	Light Emitting Diodes (LEDs)	86
2.8.5	Semiconductor Lasers	87
2.9	Processing—The Chip Level	91
2.9.1	Introduction.	91
2.9.2	Silicon Crystal Growth (1)	92

2.9.3 Epitaxy (2)	93
2.9.4 Ion Implantation (3)	94
2.9.5 Gate Oxide (4)	95
2.9.6 Polysilicon (5)	95
2.9.7 Deposited and Etched Dielectric Films (6,7,12)	95
2.9.8 Metallization (8–11)	96
2.9.9 Plasma Etching	99
2.9.10 Lithography	99
Exercises	100
References	102

Chapter 3

Defects, Contaminants and Yield	105
3.1 Scope	105
3.2 Defects in Crystalline Solids and Semiconductors	107
3.2.1 General Considerations	107
3.2.2 Point Defects	108
3.2.3 Dislocations	115
3.2.4 Grain Boundaries	122
3.2.5 Dislocation Defects in DRAMs—A Case Study	123
3.3 Processing Defects	125
3.3.1 Scope	125
3.3.2 Stress and Defects	127
3.3.3 Step Coverage	137
3.4 Contamination	141
3.4.1 Introduction	141
3.4.2 Process-Induced Contamination	142
3.4.3 Introduction to Particle Science	147
3.4.4 Combating Contamination	152
3.5 Yield	155
3.5.1 Definitions and Scope	155
3.5.2 Statistical Basis for Yield	157
3.5.3 Yield Modeling	159
3.5.4 Comparison with Experience	160
3.5.5 Yield and Reliability—What Is the Link?	162
3.5.6 Conclusion	168
Exercises	168
References	171

Chapter 4

The Mathematics of Failure and Reliability	175
4.1 Introduction	175

4.2	Statistics and Definitions	177
4.2.1	Normal Distribution Function	177
4.2.2	"Six Sigma"	180
4.2.3	Accuracy and Precision	182
4.2.4	Failure Rates	182
4.3	All About Exponential, Lognormal and Weibull Distributions	184
4.3.1	Exponential Distribution Function.	184
4.3.2	Lognormal Distribution	189
4.3.3	Weibull Distribution.	193
4.3.4	Lognormal versus Weibull	196
4.3.5	Plotting Probability Functions as Straight Lines.	196
4.3.6	Freak Behavior	199
4.4	System Reliability	202
4.4.1	Introduction.	202
4.4.2	Redundancy in a Two-Laser System	205
4.4.3	How MIL-HDBK-217 Treats System Reliability	206
4.5	On the Physical Significance of Failure Distribution Functions.	208
4.5.1	Introduction.	208
4.5.2	The Weakest Link	208
4.5.3	The Weibull Distribution: Is There a Weak Link to the Avrami Equation?	209
4.5.4	Physics of the Lognormal Distribution	211
4.5.5	Acceleration Factor	212
4.5.6	The Arrhenius Model	214
4.5.7	The Eyring Model	215
4.5.8	Is Arrhenius Erroneous?	218
4.5.9	The Bathtub Curve Revisited	219
4.6	Prediction Confidence and Assessing Risk	222
4.6.1	Introduction.	222
4.6.2	Confidence Limits	222
4.6.3	Risky Reliability Modeling.	224
4.6.4	Freak Failures	227
4.6.5	Minimizing Freak Failures	228
4.7	A Skeptical and Irreverent Summary	229
	Exercises	231
	References	235

Chapter 5

Mass Transport–Induced Failure	237
5.1 Introduction.	237

5.2	Diffusion and Atom Movements in Solids	238
5.2.1	Mathematics of Diffusion	238
5.2.2	Diffusion Coefficients and Microstructure	240
5.3	Binary Diffusion and Compound Formation	242
5.3.1	Interdiffusion and the Phase Diagram	242
5.3.2	Compound Formation	244
5.3.3	The Kirkendall Effect	245
5.3.4	The Purple Plague	246
5.4	Reactions at Metal–Semiconductor Contacts	248
5.4.1	Introduction to Contacts	248
5.4.2	Al-Si Contacts	248
5.4.3	Metal Silicide Contacts to Silicon	252
5.4.4	Contacts to GaAs Devices	253
5.5	Electromigration Physics and Damage Models	259
5.5.1	Introduction	259
5.5.2	Physical Description of Electromigration	262
5.5.3	Temperature Distribution in Powered Conductors	264
5.5.4	Role of Stress	266
5.5.5	Structural Models for Electromigration Damage	267
5.6	Electromigration in Practice	273
5.6.1	Manifestations of Electromigration Damage	273
5.6.2	Electromigration in Interconnects: Current and Temperature Dependence	276
5.6.3	Effect of Conductor Geometry and Grain Structure on Electromigration	279
5.6.4	Electromigration Lifetime Distributions	281
5.6.5	Electromigration Testing	281
5.6.6	Combating Electromigration	283
5.7	Stress Voiding	284
5.7.1	Introduction	284
5.7.2	Origin of Stress in Metallizations	285
5.7.3	Vacancies, Stresses, Voids, and Failure	288
5.7.4	Role of Creep in Stress Voiding Failure	291
5.7.5	Stress Voiding at Vias	292
5.8	Failure of Incandescent Lamps	293
	Exercises	296
	References	300

Chapter 6

Electronic Charge-Induced Damage	303
--	-----

6.1 Introduction	303
----------------------------	-----

6.2 Aspects of Conduction in Insulators	304
6.2.1 Current–Voltage Relationships	304
6.2.2 Leakage Current	307
6.2.3 SiO ₂ —Electrons, Holes, Ions, and Traps	308
6.3 Dielectric Breakdown	310
6.3.1 Introduction	310
6.3.2 A Brief History of Dielectric Breakdown Theories	311
6.3.3 Current Dielectric Breakdown Theories	312
6.3.4 Dielectric-Breakdown Testing: Ramp Voltage	316
6.3.5 Dielectric-Breakdown Testing: Constant Voltage	320
6.3.6 Ramp-Voltage and Constant-Voltage Tests: Is There a Connection?	321
6.3.7 Electric-Field and Temperature-Acceleration Factors	324
6.3.8 Plasma Charging Damage to Gate Oxides	325
6.3.9 Analogy between Dielectric and Mechanical Breakdown of Solids	326
6.3.10 Discharges and Water Trees	328
6.4 Hot-Carrier Effects	330
6.4.1 Introduction	330
6.4.2 Hot Carriers	330
6.4.3 Hot Carrier Characteristics in MOSFET Devices	331
6.4.4 Models for Hot-Carrier Degradation	331
6.4.5 Hot-Carrier Damage in Other Devices	338
6.4.6 Combating Hot-Carrier Damage	338
6.5 Electrical Overstress and Electrostatic Discharge	339
6.5.1 Introduction	339
6.5.2 ESD Models	341
6.5.3 Thermal Analysis of ESD Failures	344
6.5.4 ESD Failure Mechanisms	347
6.5.5 Latent Failures	349
6.5.6 Guarding Against ESD	351
6.5.7 Some Common Myths About ESD	351
Exercises	352
References	355

Chapter 7

Environmental Damage to Electronic Products	359
7.1 Introduction	359
7.2 Atmospheric Contamination and Moisture	360
7.2.1 Sources of Airborne Contamination	360
7.2.2 Moisture Damage	363

7.3	Corrosion of Metals	366
7.3.1	Introduction	366
7.3.2	Pourbaix Diagrams	367
7.3.3	Corrosion Rates	370
7.3.4	Damage Manifestations	370
7.4	Corrosion in Electronics	373
7.4.1	Overview	373
7.4.2	Accelerated Corrosion Testing	375
7.4.3	Corrosion of Specific Metals	376
7.4.4	Modeling Corrosion Damage	382
7.5	Metal Migration	385
7.5.1	Introduction	385
7.5.2	Metal Migration on Porous Substrates	386
7.5.3	Metal Migration on Ceramic Substrates	388
7.5.4	Mobile Ion Contamination in CMOS Circuits	390
7.6	Radiation Damage to Electronic Materials and Devices	391
7.6.1	A Historical Footnote	391
7.6.2	Some Definitions	391
7.6.3	Radiation Environments	392
7.6.4	Interaction of Radiation with Matter	393
7.6.5	Device Degradation Due to Ionizing Radiation	395
7.6.6	Soft Errors	398
7.6.7	Radiation Hardening of Devices	404
	Exercises	405
	References	408

Chapter 8

	Packaging Materials, Processes, and Stresses	411
8.1	Introduction	411
8.2	IC Chip-Packaging Processes and Effects	415
8.2.1	Scope	415
8.2.2	First-Level Electrical Interconnection	418
8.2.3	Chip Encapsulation	428
8.3	Solders and Their Reactions	435
8.3.1	Introduction	435
8.3.2	Solder Properties	438
8.3.3	Metal-Solder Reactions	438
8.4	Second-Level Packaging Technologies	447
8.4.1	Introduction	447
8.4.2	Through Hole and Surface Mounting	447
8.4.3	Ball Grid Arrays	449
8.4.4	Reflow and Wave Soldering Processes	449
8.4.5	Defects in Solder Joints	451

8.5 Thermal Stresses in Package Structures	452
8.5.1 Introduction	452
8.5.2 Chips Bonded to Substrates	454
8.5.3 Thermal Stress in Other Structures	460
Exercises	467
References	471

Chapter 9

Degradation of Contacts and Package

Interconnections	475
9.1 Introduction	475
9.2 The Nature of Contacts	476
9.2.1 Scope	476
9.2.2 Constriction Resistance	477
9.2.3 Heating Effects	479
9.2.4 Mass Transport Effects at Contacts	479
9.2.5 Contact Force	481
9.3 Degradation of Contacts and Connectors	482
9.3.1 Introduction	482
9.3.2 Mechanics of Spring Contacts	482
9.3.3 Normal Force Reduction in Contact Springs	486
9.3.4 Tribology	488
9.3.5 Fretting Wear Phenomena	489
9.3.6 Modeling Fretting Corrosion Damage	491
9.4 Creep and Fatigue of Solder	492
9.4.1 Introduction	492
9.4.2 Creep—An Overview	493
9.4.3 Constitutive Equations of Creep	494
9.4.4 Solder Creep	495
9.4.5 Fatigue—An Overview	498
9.4.6 Isothermal Fatigue of Solder	504
9.5 Reliability and Failure of Solder Joints	507
9.5.1 Introduction to Thermal Cycling Effects	507
9.5.2 Thermal-Stress Cycling of Surface-Mount Solder Joints	512
9.5.3 Acceleration Factor for Solder Fatigue	516
9.5.4 Fracture of Solder Joints	517
9.6 Dynamic Loading Effects in Electronic Equipment	526
9.6.1 Introduction	526
9.6.2 Vibration of Electronic Equipment	527
Exercises	531
References	535

*Chapter 10***Degradation and Failure of Electro-optical**

Materials and Devices	539
10.1 Introduction	539
10.2 Failure and Reliability of Lasers and Light-Emitting Diodes	540
10.2.1 A Bit of History	540
10.2.2 Reliability Testing of Lasers and LEDs	541
10.2.3 Microscopic Mechanisms of Laser Damage	547
10.2.4 Electrostatic Discharge Damage to Lasers	554
10.2.5 Contact and Bonding Reliability	556
10.3 Thermal Degradation of Lasers and Optical Components	559
10.3.1 Introduction	559
10.3.2 Thermal Analysis of Heating	560
10.3.3 Laser-Induced Damage to Optical Coatings	561
10.3.4 Thermal Damage to Lasers	562
10.4 Reliability of Optical Fibers	568
10.4.1 Introduction	568
10.4.2 Optical Fiber Strength	571
10.4.3 Static Fatigue—Crack Growth and Fracture	574
10.4.4 Environmental Degradation of Optical Fiber	578
Exercises	582
References	585

*Chapter 11***Characterization and Failure Analysis**

of Materials and Devices	587
11.1 Overview of Testing and Failure Analysis	587
11.1.1 Scope	587
11.1.2 Characterization Tools and Methods	591
11.2 Nondestructive Examination and Decapsulation	592
11.2.1 Radiography	592
11.2.2 Scanning Acoustic Microscopy	594
11.2.3 Analysis of PIND Particles	598
11.2.4 Decapsulation	599
11.3 Structural Characterization	603
11.3.1 Optical Microscopy	603
11.3.2 Electron Microscopy	604
11.3.3 Transmission Electron Microscopy	611
11.3.4 Focused Ion Beams	612

11.4 Chemical Characterization	615
11.4.1 Introduction	615
11.4.2 Making Use of Core-Electron Transitions	615
11.4.3 Chemical Analysis by Means of Ions	619
11.5 Examining Devices Under Electrical Stress	621
11.5.1 Introduction	621
11.5.2 Emission Microscopy	621
11.5.3 Voltage Contrast Techniques	625
11.5.4 Thermography	630
11.5.5 Trends in Failure Analysis	635
Exercises	635
References	638

Chapter 12

Future Directions and Reliability Issues	641
12.1 Introduction	641
12.2 Integrated Circuit Technology Trends	642
12.2.1 Introduction	642
12.2.2 IC Chip Trends	642
12.2.3 Contamination Trends	648
12.2.4 Lithography	650
12.2.5 Packaging	651
12.3 Scaling	654
12.3.1 Introduction	654
12.3.2 Implications of Device Scaling	656
12.4 Fundamental Limits	658
12.4.1 Introduction	658
12.4.2 Physical Limits	659
12.4.3 Material Limits	660
12.4.4 Device Limits	660
12.5 Improving Reliability	662
12.5.1 Reliability Growth	663
12.5.2 Failure Prediction: Stressor-Susceptibility Interactions	664
12.5.3 Building In Reliability	665
Exercises	669
References	671
Appendix	673
Index	675

Chapter 2

Electronic Devices: How They Operate and Are Fabricated

2.1 INTRODUCTION

Books on reliability ideally stress issues related to component and device failure within the context of how they are intended to function. Because it is often not necessary to know about the latter to address the former, these two bodies of knowledge and expertise often remain separate concerns. This is especially so because most failures involve materials other than the semiconductors, e.g., metallizations, insulation, and packaging materials. Nevertheless, it is fair to say that malfunction, failure, and reliability can be more intelligently confronted through some knowledge of device behavior and fabrication methods employed. For example, awareness of device current-voltage characteristics is certainly helpful in diagnosing failure modes in some instances. Similarly, an appreciation of the processing steps involved in device manufacture may suggest ways to avoid future failures. In this spirit, the chapter is primarily devoted to elevating the consciousness and sensitivity of readers to electronic devices and their uses, characteristics, and instabilities. Many aspects of the degradation and reliability of specific devices will be introduced here in a qualitative way. We shall return to these same issues in later chapters, where degradation and the physics of failure will be analytically modeled at a deeper level.

A broad tutorial survey of electronic materials initiates the chapter, which then goes on to treat the operation of diodes, transistors, and optical devices. For the most part, the silicon and compound semiconductor devices introduced are those that have been used most broadly in practice. Instead of mathematical rigor in describing device behavior, equations will be used sparingly and introduced without derivation. Qualitative and intuitive meanings will be emphasized. In doing so it is hoped that the language and terminology used to describe devices will make references to them more understandable later in the book as well as in the reliability literature. Following this, several of the processing steps involved in device manufacture are introduced. Defects and processing problems, the probable sources of a number of reliability concerns, are discussed in the next chapter.

Electrical engineers will probably choose to skip much of this chapter because of its admittedly shallow coverage; however, the latter part dealing with processing may cover new ground. Conversely, the nomenclature and qualitative descriptions of electronic materials and devices may prove instructive to those trained in other disciplines. In either case, several excellent books have been written on the physics of solid state electronic devices (1–5) and on processing technology (6–10) that all readers can profit from; I certainly have.

2.2 ELECTRONIC MATERIALS

2.2.1 Introduction

At the outset it is important to realize that a very large number of materials have been employed at one time or another in experimental and commercial electronic devices and equipment. They form the sum and substance of active devices and passive components, the packages that enclose them, the substrates and boards on which they are mounted, and the solder and wires that join and connect them. All commercial electronic materials and devices have typically survived a rigorous selection ordeal whose hurdles included:

1. meeting stringent electrical specifications and requirements,
2. developing reproducible processing and achieving high yields in manufacturing,
3. extensive testing to assure reliable operation.

We are concerned with the properties of these surviving materials and devices and their fates during subsequent use. The survivors, however, are typically but a fraction of the total number of potential candidates that were initially consid-

ered. One only has to recall Edison's experimentation with lamp filament materials to appreciate what is involved.

A list of the materials employed in commercial electronic devices is entered in Tables 2-1 and 2-2. Although incomplete, the list includes most of the materials addressed in this book. These materials can be conveniently classified into three categories, namely, *conductors*, *semiconductors*, and *insulators*. Each category is distinguished both by the magnitude of the electrical resistivity as indicated in Fig. 2-1 and by the sign of the temperature coefficient of resistivity. By *conductors* we mean metals with resistivities ranging from 10^{-6} to perhaps 10^{-4} ohm-cm. Semiconductor resistivities depend very strongly on doping level and span a typical range from 10^{-3} to 10^6 ohm-cm. However, extremely heavily doped Si has metallic resistivities, while semi-insulating GaAs has a resistivity of $\sim 10^8$ ohm-cm. Insulator resistivities are the highest of all and can reach values in excess of 10^{15} ohm-cm. In electronic applications, insulators are represented by oxides, e.g., SiO_2 in IC chips, Al_2O_3 substrates on which devices are mounted, and polymers, e.g., IC encapsulants and printed circuit boards.

Physically, metals can be distinguished from all other classes of materials because they become *poorer* conductors as the temperature is raised, implying a positive temperature coefficient of resistivity. Conversely, semiconductors and insulators become *better* conductors as they become hotter. Even here there are rare exceptions to the rule over small temperature excursions.

2.2.2 Semiconductors

2.2.2.1 Structure Semiconductors and the devices based on them are a primary focus of the book. Some, but certainly by no means all, of the reasons for device failure and reliability problems can be attributed to malfunction of semiconductors. The special attributes of semiconductors derive from both their physical and electronic structures. In Fig. 2-2 the crystal structures of both Si and GaAs, the two most important semiconductors used in electronics, are depicted. Silicon has a diamond-cubic structure that is derived from the face-centered cubic lattice but with a two-atom motif per lattice point; this means that there are eight atoms per unit cell. On the other hand, in GaAs the Ga and As atoms alternate as shown within the basic diamond cubic lattice.

In cubic structures there are three important crystallographic planes and directions. The (100) planes are the cube faces and the [100] directions are the cube edges. Planes passing through diagonally opposite cube edges have (110) indices, while (111) planes pass through 3 diagonally opposite cube corners. The [110] directions are those between the cube center to the midpoint of each of the 12 cube edges; similarly, [111] directions radiate from the cube center to each of

Table 2-1 Physical Properties of Electronic Materials

Metals	Melting point (°C)	Electrical resistivity (Ω -cm)	Temp. coeff. resistivity (α) (°C ⁻¹)	Density (g/cm ³)
Copper	1083	1.7×10^{-6}	0.0039	8.93
Silver	962	1.6×10^{-6}	0.0041	10.5
Gold	1063	2.2×10^{-6}	0.0034	19.3
Aluminum	660	2.7×10^{-6}	0.0039	2.7
Tungsten	3415	5.5×10^{-6}	0.0045	19.3
Platinum	1772	10.6×10^{-6}	0.0039	21.5
Palladium	1552	10.8×10^{-6}	0.0038	12.0
Nickel	1455	6.8×10^{-6}	0.0069	8.9
Chromium	1875	13×10^{-6}	0.003	7.2
Invar (36Ni-64Fe)	1425	80×10^{-6}	0.0014	8.0
Alloy 42 (43Ni-58Fe)	1425	68×10^{-6}	0.0014	8.1
Kovar	1450	50×10^{-6}		8.3
63Sn-37Pb	183	15×10^{-6}		8.42
95Pb-5Sn	310	19×10^{-6}		11
Cu-W (20%Cu)	1083	2.5×10^{-6}		
CoSi ₂	550*	15×10^{-6}		
TiSi ₂ (C49)	650*	60×10^{-6}		
TiSi ₂ (C54)	870*	15×10^{-6}		
WSi ₂	650*	100×10^{-6}		
PtSi	300*	35×10^{-6}		

the 8 cube vertices. Every semiconductor electronic device requires either bulk or thin film semiconductor single crystals of high crystallographic perfection.

Critical to many of the structural and electronic properties of elemental Si is the tetrahedral atomic grouping, in which each atom is surrounded by four identical atoms. The bonds between atoms, represented by lines, are covalent in nature, reflecting the fact that a pair of outer electrons, one from each of the neighboring atoms, comprises the bond. Covalent bonds are typically strong and resistant to deformation. This accounts for silicon's appreciable hardness, relatively high Young's modulus, lack of ductility, and high melting point. Although covalent bonds are mechanically strong, the bonding, or *valence*, electrons within them can be easily detached at relatively low temperatures to become more energetic *conduction* electrons. Holes, with an effective positive charge, are then left behind in the broken valence bonds.

Both holes and conduction electrons can readily move under the influence of electric fields. The semiconductor property known as mobility (μ) is defined as the ratio of the carrier velocity to the applied electric field and has units of

TABLE 2-1 continued

Ceramics/ Glasses	Melting point (°C)	Electrical resistivity (Ω -cm)	Dielectric constant	Thermal conductivity (W/m-K)	Density (g/cm ³)
Silicon dioxide	1980	$>10 \times 10^{15}$	3.8	7	2.2
Aluminum oxide	2323	$>10 \times 10^{14}$	9.9	37	3.9
Silicon nitride	2173	$>10 \times 10^{14}$	7.0	30	3.2
Silicon carbide	3110	$>10 \times 10^{14}$	42	270	3.2
Aluminum nitride		$>10 \times 10^{11}$	8.8	140–230	3.3
Beryllium oxide	2570	$>10 \times 10^{14}$	6.8	240	2.9
Boron nitride		$>10 \times 10^{10}$	6.5	600	1.9
Borosilicate glass	1075	$>10 \times 10^{10}$	4.1	5	
Plastics					
Epoxy-Kevlar	200*	$>10 \times 10^{14}$	3.6	0.2	
Polyimide	200*	$>10 \times 10^{14}$	4.0	0.35	
Epoxy FR-4	175*	$>10 \times 10^{14}$	4.7	0.2	1.9
Polyimide	300–400*	$>10 \times 10^{15}$	3.0–3.5	0.2	1.5
Teflon	400*	$>10 \times 10^{14}$	2.2	0.1	2.2
Silicone (high-purity)	150*	$>10 \times 10^{14}$	2.7–3.0	0.16	
Parylene N (Vapor deposited)		$>10 \times 10^{14}$	2.65	0.12S	

*Processing temperature

Note: Additional properties are listed in Table 8-2.

Data collected from various sources.

(cm²/V-s). Mobility values for carriers in semiconductors are listed in Table 2-2. Undesirable reductions in carrier mobility occur in the presence of impurities and lattice defects. The combination of carrier concentrations and velocities influences and limits all electrical conduction phenomena.

2.2.2.2 Bonds and Bands

a. Bonds Several representations of electronic structure have evolved over the years as aids in interpreting the behavior of semiconductor materials and devices. The simplest of these, depicted in Fig. 2-3, is a physical-bonding model where the tetrahedra are flattened to two dimensions, and each bond between atoms now represents a couple of shared valence electrons. Introduction of n-type dopants (P, As, and Sb) and p-type dopants (B, In, and Al) into the Si lattice creates extra donor electrons and extra acceptor holes, respectively. Through doping, the number of electron and hole charge carriers can be dramatically

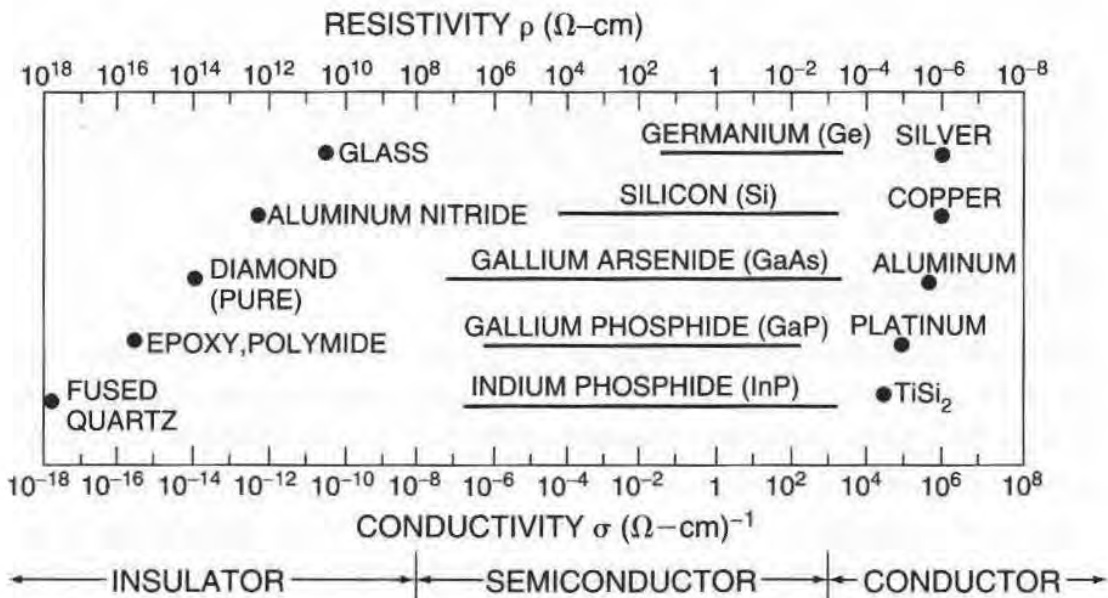
Table 2-2 Properties of Elemental and Compound Semiconductors

Material	Lattice Parameter (nm)	Melting Point (K)	Energy Gap (eV at 25°C)	Electron Mobility (cm ² /V-s)	Hole Mobility (cm ² /V-s)
Diamond	0.3560	~4300	5.4	1800	1400
Si	0.5431	1685	1.12 I	1450	450
Ge	0.5657	1231	0.68 I	3600	1900
ZnS	0.5409	3200	3.54 D	120	5
ZnSe	0.5669	1790	2.58 D	530	28
CdS	0.5832	1750	2.42 D	340	50
AlAs	0.5661	1870	2.16 I	1200	420
GaP	0.5451	1750	2.26 I	110	75
GaAs	0.5653	1510	1.43 D	8500	400
InP	0.5869	1338	1.27 D	4600	150
InAs	0.6068	1215	0.36 D	30000	460

I refers to indirect band gap.

D refers to direct band gap.

From S.M. Sze, *Semiconductor Devices—Physics and Technology*, J. Wiley & Sons, NY (1985), and B. R. Pamplin in *Handbook of Chemistry and Physics*, Ed. R. C. Weast, CRC Press, Boca Raton FL (1980).

**Fig. 2-1** Electrical resistivities of electronic materials.

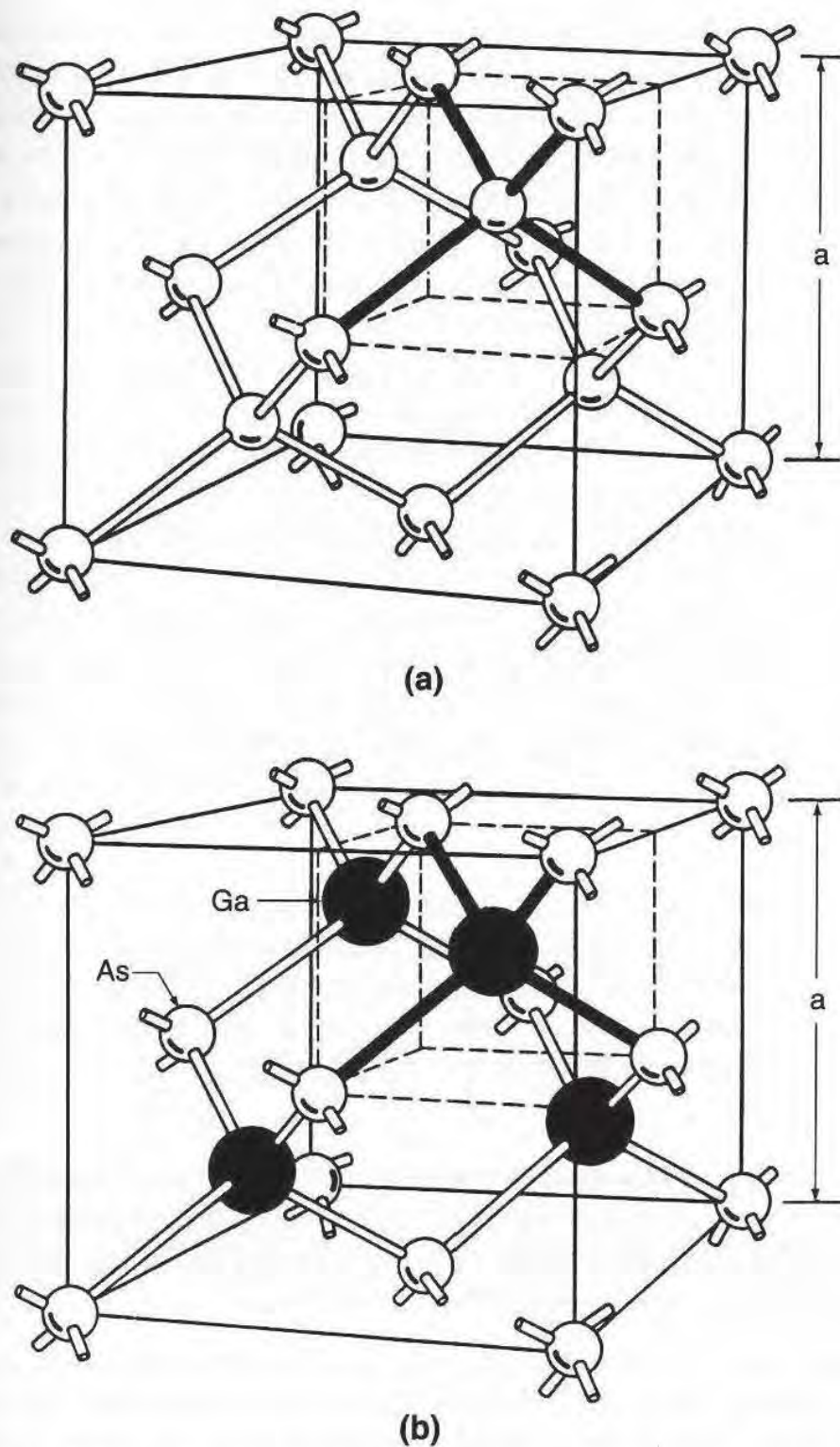


Fig. 2-2 (a) Diamond cubic crystal structure of silicon. (b) Model of the zincblende crystal structure of GaAs. Both Ga and As populate face-centered cubic lattices that interpenetrate one another.

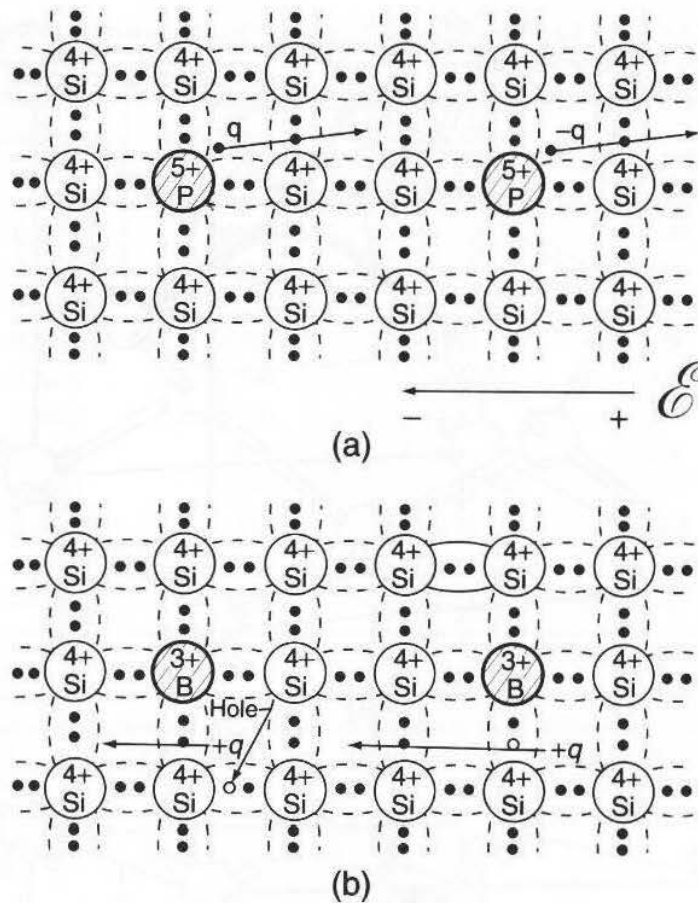


Fig. 2-3 (a) Representation of a two dimensional silicon lattice containing donor (P^{+5}) atoms. (b) Two-dimensional silicon lattice containing acceptor (B^{+3}) atoms. Each bond consists of two valence electrons. Direction of carrier motion in electric field $\mathcal{E}$ is shown.

enlarged in a controllable way. In this model, conduction electrons hop around in the open spaces between bonds; holes, on the other hand, are confined to, and migrate through, the bond network. An applied electric field causes these charge carriers to move in opposite directions.

b. Bands The electronic structure of undoped as well as doped semiconductors is popularly represented in terms of the energy band diagrams of Fig. 2-4. When isolated atoms of Si are brought together to form the solid, individual electron energy levels collectively broaden to create bands containing an astounding 10^{22} or so distinct levels (one for each electron) per cubic centimeter. They are packed so tightly that the energy difference between the highest and lowest levels is only a few electron volts. Thus all of the electrons comprising valence bonds form the *valence band*, and all of the conduction electrons enter

the *conduction band*. The energies of electron levels within states are noted on the ordinate, while the horizontal axis can be loosely identified with some macroscopic lattice distance having much larger than atomic dimensions. A key feature of this diagram is the *energy gap* (E_g) that separates valence and conduction bands. In Si, $E_g = 1.12$ eV, while in GaAs, $E_g = 1.43$ eV at 300K. Electron energy levels and states within the energy gap are forbidden in pure, structurally perfect Si. Any perturbation (e.g., dopant atom, impurity located either on lattice or interstitial sites, vacancy or dislocation defects, precipitates) that disturbs this ideal order is reflected in the creation of new quantum states with characteristic energy levels. Such states usually fall somewhere in the gap.

On this basis it is easy to see that filled n-type dopant donor states (Fig. 2-4a) lie close to the bottom of the conduction band, to which they readily donate electrons. Similarly, empty p-type dopant acceptor states (Fig. 2-4b) lie just above the top of the valence band, where they accept electrons, leaving holes behind in the process. In essence, this is how electrical conduction occurs in semiconductors. An energy level of interest that reflects the probability of carrier population of states is known as the Fermi (E_F) level. It is defined to be the energy level whose probability of occupation is 0.5. In n- and p-type semiconductors, E_F lies closer to either the conduction (E_C) or valence (E_V) band edges, respectively, but usually within the gap.

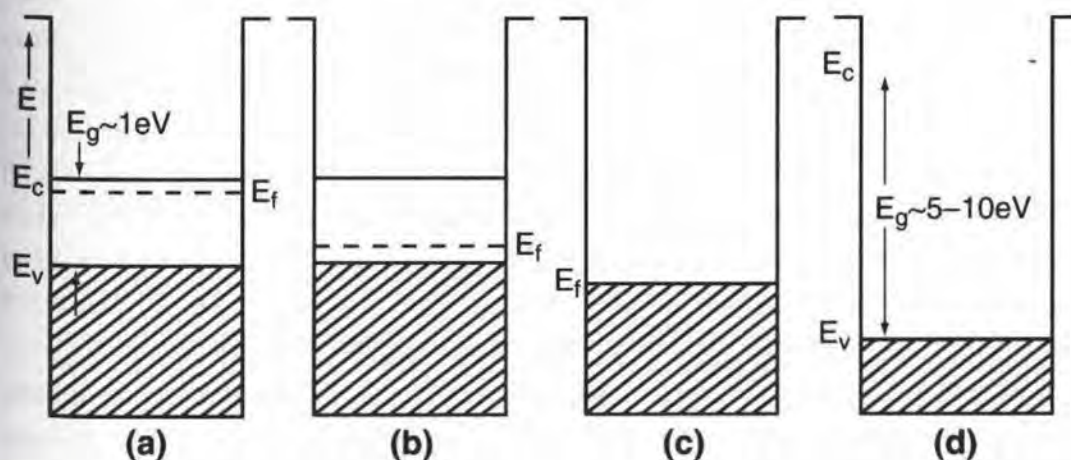


Fig. 2-4 Electron energy band diagrams for semiconductors, metals, and insulators. (a) n-type semiconductor. (b) p-type semiconductor. Largely filled valence band lies below E_V and largely empty conduction band lies above E_C . The Fermi level lies closer to E_C in the n-type semiconductor and closer to E_V in the p-type semiconductor. (c) Band diagram of a metal. The uppermost occupied electron state in the distribution corresponds to the Fermi level, above which there are always accessible empty electron states. There is no energy gap in metals. (d) Band diagram for insulators. Electron energy gaps range from 5 to 10 eV.

Band diagrams are helpful in distinguishing the behavior of semiconductors from those of metals and insulators. In metals (Fig. 2-4c) the Fermi level caps the electron distribution. Below it all levels are occupied by electrons, and above it they are unoccupied. It takes very little energy to access these closely spaced empty levels if the electrons are stimulated electrically, thermally, and optically; this accounts for the excellent electrical and thermal conductivity as well as high reflectivity of metals. Band diagrams for insulators (Fig. 2-4d) resemble those for semiconductors, except for the larger electron energy gaps ranging from 5 to 10 eV.

c. Energy–Momentum Representation Another widely used representation of semiconductor bands is shown in Figs. 2-5a and b. Here the energy (E) of electron states is plotted as a function of the magnitude and direction of the

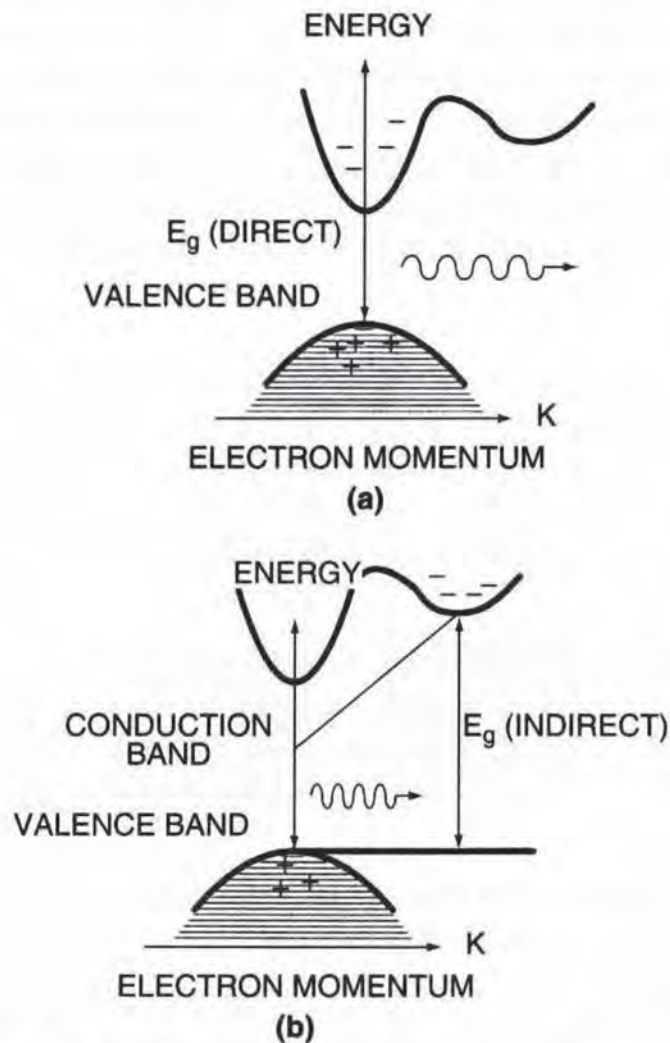


Fig. 2-5 Energy (E) versus wave vector (k) representation of band structure. (a) direct band-gap semiconductor. (b) indirect band-gap semiconductor.

electron momentum, or equivalently, versus wave vector (k). The resultant parabola-like E vs k curve is important because it provides a useful way to distinguish between Si or Ge on the one hand, and GaAs as well as similar compound semiconductors on the other (1). While the former are known as *indirect band-gap* semiconductors, the latter are *direct band-gap* semiconductors. Only direct-gap semiconductors are useful in optical devices like light-emitting diodes (LED) or lasers. Photons of the band-gap energy are emitted when electrons descend from the bottom of the conduction band to recombine with holes at the top of the valence band. The transition between the two involved states must conserve electron momentum. Similarly, when a photon whose energy is equal to or greater than the gap energy is absorbed, momentum conservation governs the electron ascent between the corresponding valence and conduction band states. In direct band-gap materials like GaAs and InP these electron descents and ascents are vertical, conserving momentum; but in Si and other indirect semiconductors the transitions are “slanted” and involve momentum changes that effectively excite lattice vibrations which heat the solid. These wasteful processes dramatically reduce the probability that photons will either be emitted or absorbed during respective electron-hole recombination or generation.

2.2.2.3 Compound Semiconductors The important III–V compound semiconductors are derived from atomic combinations arising from columns III A (e.g., Al, Ga, In) and V A (e.g., N, P, As) of the periodic table. Other II–VI compound semiconductors are composed of elements from columns II B (Zn, Cd, Hg) and VI A (S, Se, Te). A summary of the electrical properties of elemental and compound semiconductors is reproduced in Table 2-2, where lattice parameters, energy gaps, and mobilities are indicated for the more widely used materials. Ternary (3 atom) and quaternary (4 atom) semiconductors are also widely used in electro-optical applications. For example, AlGaAs ternaries are derived from AlAs and GaAs mixtures. Similarly, InGaAs and GaInAsP devices used in optical communications can be understood as combinations of binary compounds, e.g., GaAs and InAs, and GaAs and InP. In many instances the resultant ternary energy gap and lattice parameter values are simply weighted averages derived from the constituent binaries. Thus by alloying GaAs and GaP, families of light-emitting diodes with compositions $\text{GaAs}_{1-x}\text{P}_x$ become available, provided that the ternary semiconductor alloys are still direct-gap materials. For $x = 0.4$, interpolation yields an energy gap of ~ 1.9 eV, which from Eq. 2-1 below corresponds to the red portion of the spectrum.

When films of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ deposited on GaAs substrates, or InGaAs deposited on InP, are involved, it is essential that the film be epitaxial. This means that film and substrate must be “lattice matched” or have as nearly identical lattice para-

meters as possible. Otherwise, defects that severely limit device function will tend to be present at the film–substrate interface. Typically, mismatches in lattice parameter of no more than 0.1% can be tolerated.

2.2.2.4 Electrical and Optical Function Semiconductors have several unique properties that enable them to exhibit unusual electrical and optical effects. Like metallic conductors they possess sufficient numbers of high-mobility charge carriers to pass large currents. Unlike conductors, but like insulators, they are capable of supporting large electric fields; this happens at p–n junctions under reverse bias where charge carriers are depleted. Remarkably, in semiconductor devices either attribute can be emphasized in applications requiring rectification (diode) and amplification (transistor) of current, or storage of charge (MOS capacitor). Application of the requisite bias voltages to the device serves to control the magnitude of current flow.

Optical activity is possible in semiconductors because they possess an important pair of electron levels that spans the energy gap. Many states of very nearly the same energy are either full or empty at levels near the bottom of the conduction band and the top of the valence band; thus electron transitions can readily occur. The wavelength (λ) associated with the band-gap energy is derived from the well known Planck equation, $E = hc/\lambda$ (i.e., $E_g = hc/\lambda$), where h is Planck's constant and c is the speed of light. A useful formula connecting E_g and λ is

$$\lambda \text{ (microns)} = 1.24/E_g \text{ (eV)}. \quad (2-1)$$

Three basic types of electro-optical phenomena can be simply understood by referring to Fig. 2-6. Depicted are the valence and conduction bands of the involved semiconductor or active electro-optical medium, and the radiative transitions that occur between energy levels in them. Either photons or electronic signals can trigger mutual electro-optical responses at p–n junctions.

1. *Stimulated absorption* of incident photons (Fig. 2-6a) excites electrons from the valence to the conduction band, generating mobile carriers or an electric signal. After photon detection (e.g., by a photodiode) electron circuitry is activated to amplify the signal, or deliver external power (e.g., a solar cell).
2. The reverse process of photon emission occurs during deexcitation of electrons from the conduction-to-valence band (Fig. 2-6b). In light-emitting diodes (LED), for example, the emission is triggered by an electric current that passes through the junction causing electron–hole recombination and a corresponding *spontaneous emission* of monochromatic light. The light can now be viewed directly in a display, or made to pass through lenses of an optical system, or enter an optical fiber.

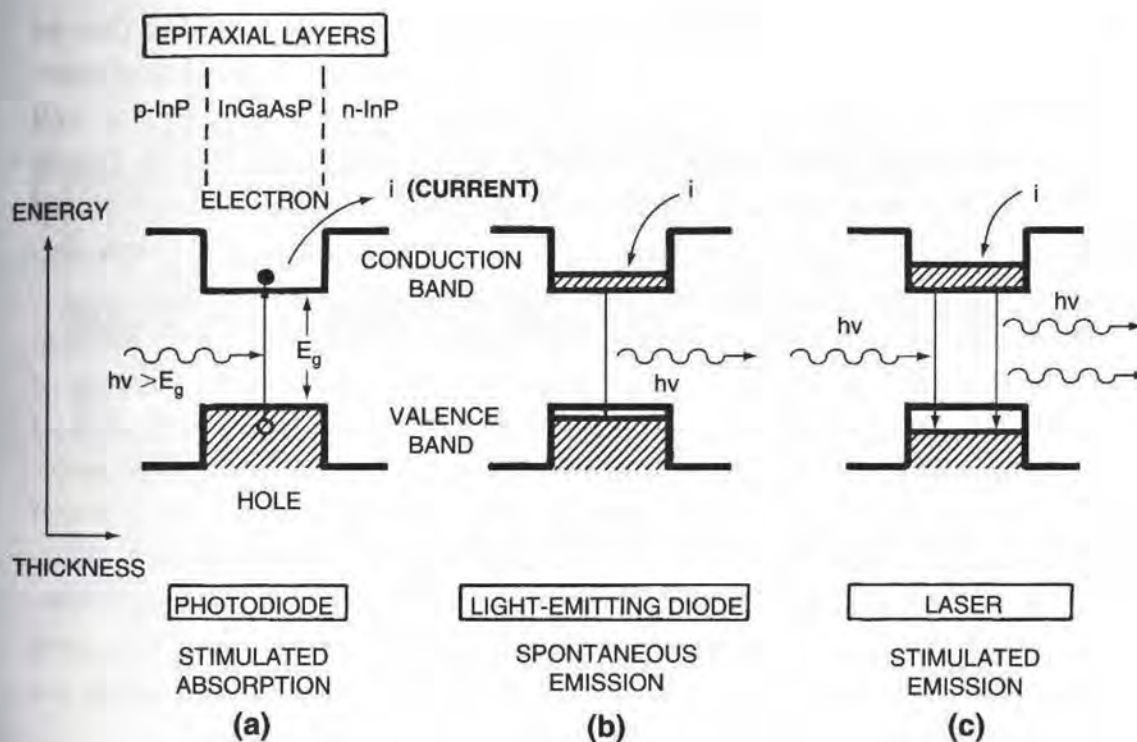


Fig. 2-6 Basic electron transitions between valence and conduction bands in semiconductors shown within the context of junction devices. (a) Stimulated absorption. (b) Spontaneous emission. (c) Stimulated emission.

- The last case (Fig. 2-6c) combines features of the previous two. Through absorption of light (pumping), electrons ascend from the valence band to an excited state, causing a normally unnatural, inverted population of electron levels. Incident photons that impinge on such excited atoms may now induce *stimulated emission* of photons from them. These photons have the same energy and phase as those that excited them in the first place. If instead of allowing the light to exit the medium it is made to reflect back and forth in phase, while impinging on other atoms in excited states, the light intensity is amplified. This is the basis of laser action.

Just how the electro-optical effects are capitalized upon in actual devices is addressed later in the chapter.

2.2.3 Conductors

Conductors are used in circuits to speed the transmission of electronic signals from where they are generated to where they are processed, e.g., detected, amplified, stored, or displayed, without distortion or any loss of information. In solid-state circuits this means that conductors must first make low-resistance contact

to various n and p doped regions, and gate electrodes of devices, and then be connected to other active devices and passive components via a network of interconnections. The latter are found in great profusion on the IC chip itself as well as at the (packaging) interface between the chip and the outside world. Degradation of both contacts and interconnections may alter the fidelity of the intended signal by changing its amplitude, time constant, sign, and shape; frequently, open circuiting is the final manifestation of conductor degradation.

Properties required of conductors include high electrical conductivity (to minimize both Joule heating and RC time constants), chemical stability (absence of corrosion), ability to be easily processed into required shapes with desired dimensions, compatibility with other device processing and fabrication steps, and lastly, resistance to loss of dimensional and physical integrity due to stress or the passage of large current densities (electromigration). Because the conductivity requirement is often the critical one, conductors are limited to metals. These materials have face-centered and body-centered cubic crystal structures for the most part. Unlike semiconductors, in all practical applications metals are polycrystalline. Relevant properties of metals are given in Table 2-1.

Of the 80 or so metallic elements only a few (e.g., Al, Cu, Au, Ag) have seen extensive use in commercial electronic devices and systems. On IC chips, interconnections are basically limited to Al and Al alloy (Al-Cu, Al-Cu-Si) metalizations at any given level, and to W/TiN via combinations to connect levels. But, when contacts to semiconductors and external connectors are considered, a much richer assortment of metals, alloys (Ti-W, PdSi, Sn-Pb solder), and intermetallic compounds (TiSi_2) is utilized. Metal contacts and interconnections at several levels within ICs are schematically indicated in Fig. 2-7 together with typical sites and types of degradation.

Metals are good conductors of electricity because they contain huge numbers of free or conduction electrons; i.e., Cu has 1 and Al has 3 per atom, so we are dealing with $\sim 10^{22}$ per cm^3 . Their band structures are complex, but importantly, metals possess either overlapping valence and conduction bands or an unfilled conduction band. Further, the absence of energy gaps means that electrons never face any appreciable barrier in accessing the numerous, closely spaced, empty higher energy states under the stimulus of an applied electric field or incident photon. Two simple relationships govern the resistivity, ρ , of metals:

$$\rho_{\text{Total}} = \rho_T + \rho_i + \rho_d \quad (2-2)$$

and

$$\rho = \rho_0 (1 + \alpha \Delta T). \quad (2-3)$$

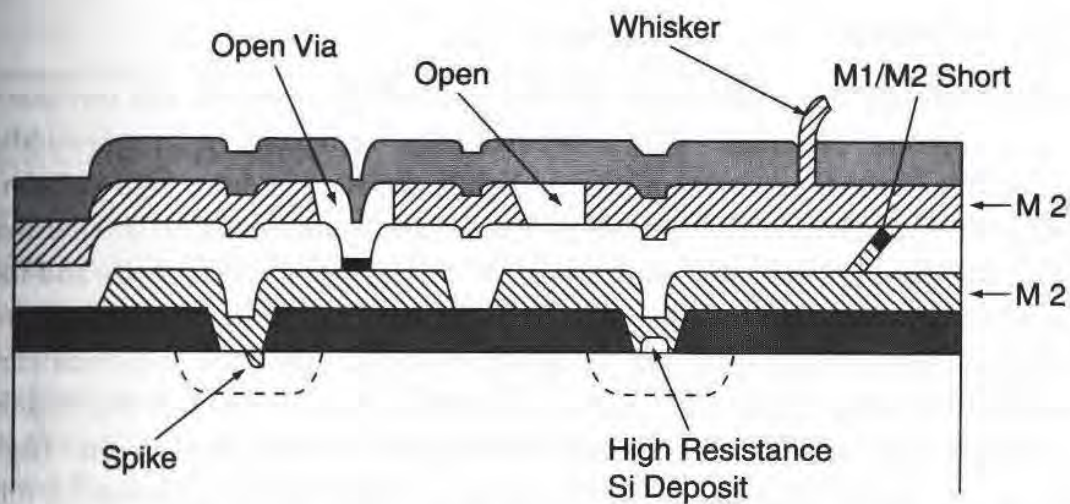


Fig. 2-7 A schematic of the variety of metal contacts and interconnections in integrated circuits together with typical degradation phenomena and sites of occurrence.

The first is a statement of Matthiessen's rule, which reflects the fact that the total resistivity of a metal has three contributions:

1. a thermal one, ρ_T , which increases linearly with temperature in the vicinity of 25°C. Since ρ_T generally dominates the other two contributions in this temperature range, Eq. 2-3 describes the observed resistivity (ρ) of metals for all practical purposes. In this formula ρ_0 is the resistivity at 25°C, α is the temperature coefficient of resistivity, and ΔT is the difference in operating temperature relative to 25°C.
2. ρ_i , the resistivity due to atomic impurities (dissolved or in compound form).
3. ρ_d , the resistivity due to defects.

Alloying always increases the resistivity of the base metal. But compared to orders of magnitude change in semiconductor resistivity with minute doping levels, metals suffer relatively little conductivity loss with much larger alloying additions. Lattice defects such as vacancies, dislocations, and grain boundaries raise the resistivity of metals, but only slightly. When impurity or defect concentrations are small, their effect on the total resistivity can be measured only at cryogenic temperatures, where the thermal contribution (ρ_T) is very small by comparison, i.e., frozen out. During arcing or Joule heating the conductivity of metals may degrade due to oxidation or interdiffusion reactions that cause extensive compositional change. Interconnections suffering electromigration damage invariably show an increase in resistance.

2.2.4 Insulators

We have already noted that in the scheme of conductors, metals and insulators are at opposite extremes, with semiconductors occupying a broad middle ground. When the latter are very pure or lightly doped, they have high resistivities, and their dielectric character is emphasized. Interestingly, similarities between semiconductors and insulators exist. Energy gaps exist in both, and the response of conductivity to doping and temperature is similar. Insulators are used to isolate conducting regions at different electric potentials so that no current flows between them. They also serve vital dielectric functions that enable circuits to store charge in capacitors that behave as memory elements. High dielectric constant ferroelectric ceramics such as BaTiO_3 are important in this regard for potential inclusion in memory chips. Dielectric property change upon cyclic electrical stressing is a concern, however.

The insulating materials we are primarily concerned with are oxides (e.g., SiO_2 , Al_2O_3) and polymers. Invariably these insulators have amorphous structures. Interestingly, solid insulators are not the only ones of interest in electrical equipment. Transformer oil is an example, and its breakdown is an important consideration in the reliability of very high voltage equipment. On the IC chip level, thin films and coatings of silicon dioxide, silicon nitride, and silica-based glasses are used in assorted insulating, dielectric, processing, and device passivation roles. These film materials are deposited at relatively low temperatures by both physical and chemical vapor deposition techniques. The most important insulator in this book is SiO_2 because it plays critical roles in two important diverse electronic technologies, for quite different reasons. Not only does SiO_2 serve as the thin film gate insulator in field effect transistors, but in fiber form it is the medium through which light propagates in optical communications systems. In these applications it has been the focus of dielectric as well as mechanical breakdown investigations that will be respectively discussed in Chapters 6 and 10.

Since insulators have very few electron and hole carriers, conduction by these carriers is limited, a fact due to the large value of E_g . Thin oxide and glass films are actually quite a bit more conductive than bulk insulators. For one thing there are fuzzy electron-distribution tails at both the top of the valence band and the bottom of the conduction band. These extend into the normally state-free gap and overlap to create a broad continuum of electron states. Secondly, the electrode contacts influence the nature of electron injection into the insulator. Lastly, because dimensions are small, modest applied voltages all too often lead to large electric fields that, in turn, may foster nonlinear conduction behavior, charge trapping, and sometimes breakdown. Later, in Chapter 6, other conduction mechanisms will be introduced. Questions of how electrons interact with matrix atoms and other electrons in the insulator to generate damage will be addressed then.

2.3 DIODES

2.3.1 The p–n Junction

The fundamental building block of most semiconductor devices is the p–n junction, i.e., p and n doped regions joined at a junction. Diodes contain one such junction, transistors two, and there are devices that have three or more junctions. Understanding the operations of rectification, amplification, and switching implies some knowledge of junction behavior. Immediately after p and n doped regions come into contact or are shorted together, holes and electrons rapidly traverse the junction and, as shown in Fig. 2-8a, three things happen.

1. Carrier equilibrium results in equalization of the respective Fermi levels.
2. A built-in electric contact potential (V_0) or field ($\mathcal{E}$) is established that discourages any further charge transport.
3. In effect, a space charge or depletion region (a kind of dielectric, carrier-free zone) develops.

The extent of the space charge region varies inversely with doping level. When the latter is high, the material is highly conductive and cannot support high electric fields. Just the opposite is true at low doping levels. Thus, devices that function at low voltages are generally more highly doped than those exposed to high voltage levels.

Importantly for device behavior, the internal barrier to charge flow can be manipulated by applying an external bias voltage. When the diode is reverse biased by connecting n to the positive supply voltage (V) and p to the negative supply voltage (Fig. 2-8b), all of the p-type Si electron levels, including the Fermi level, rigidly rise relative to the n-type Si levels by an amount qV . Electrons and holes are drawn away from the junction in the respective n and p regions. We speak of a *depletion* region free of carriers. This has the effect of widening the space charge region as well as raising the magnitude of the potential barrier and electric field across the junction. As a result, only a very small reverse current (i_R) flows. Alternatively, when the diode is forward biased by connecting n to the negative supply voltage and p to the positive supply voltage (Fig. 2-8c), electron levels in the two semiconductors now rigidly shift in the other direction relative to each other. Electrons and holes now readily flow across the junction where the carriers recombine. This has the effect of reducing the space charge region as well as lowering the magnitude of the potential barrier and electric field across the junction. Importantly, electron–hole recombination means a large forward current (i). The p–n junction thus enables currents of vastly differing magnitude to flow by simply changing the bias voltage polarity.

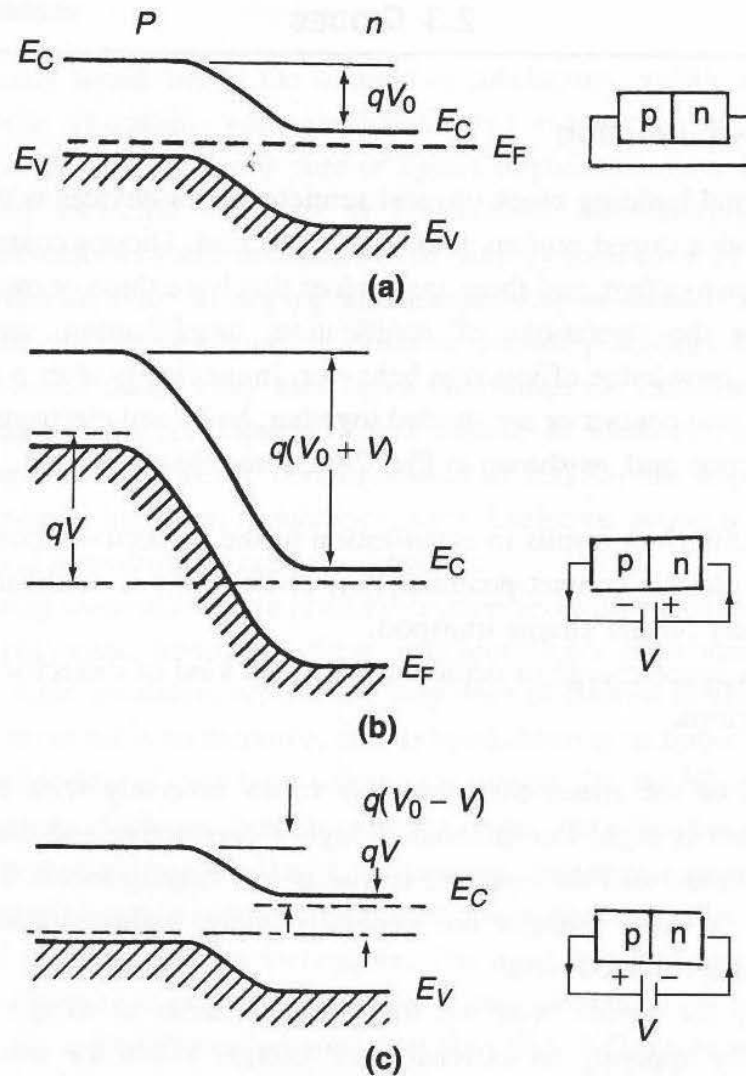


Fig. 2-8 Energy band representation of a semiconductor p-n junction under different conditions of bias. (a) Zero bias voltage—The Fermi levels are equilibrated on both sides of the junction, creating an internal electric field. (b) Reverse bias—The Fermi levels are displaced as carriers are depleted from the junction region. A small reverse current flows. (c) Forward bias—Large forward current flows as carriers recombine at the junction.

This is the behavior desired of rectifiers. In contrast, ohmic characteristics are linear and display the same current magnitudes, but of opposite sign, when the voltage polarity reverses.

The ideal junction characteristics, depicted in Fig. 2-9, are succinctly governed by an equation that describes the current-voltage variation as

$$i = i_R \left[\exp\left(\frac{qV}{kT}\right) - 1 \right], \quad (2-4)$$

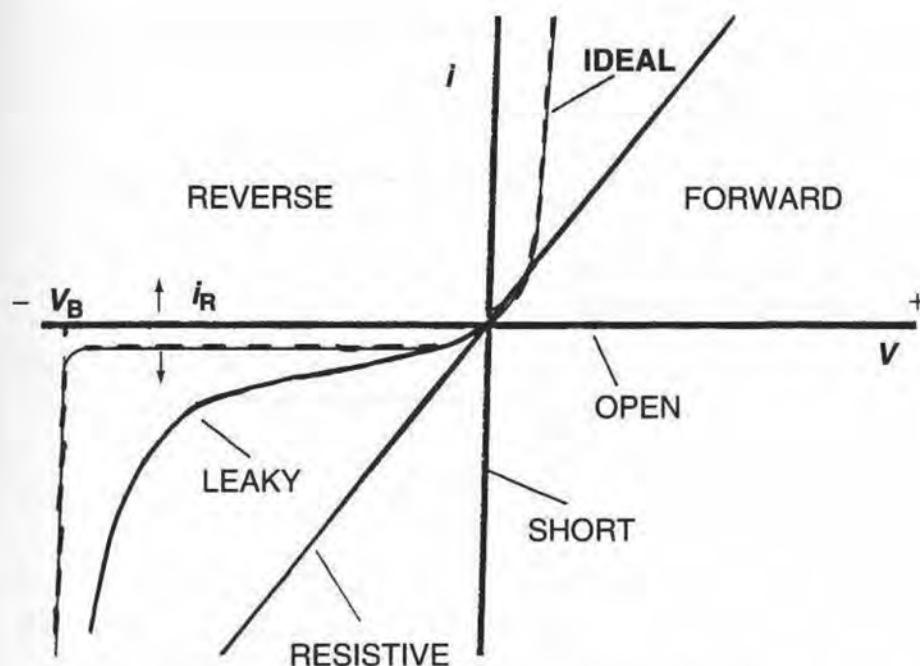


Fig. 2-9 Ideal and nonideal i - V junction characteristics.

where q is the electronic charge, and kT has the usual meaning. Although dc operation has been described, junctions in devices generally function under ac transient conditions. This causes voltages, currents, and internal fields to be time-dependent.

2.3.2 Contacts

When metal contacts are applied to the p and n semiconductor regions, low-resistance ohmic behavior is normally desired. The band diagram for an idealized ohmic contact is shown in Fig. 2-10a. In practice, however, ohmic contacts are prepared by heavily doping the underlying semiconductor in order to minimize its resistivity. This considerably narrows the region of band bending, promoting bidirectional tunneling of charge across the metal-semiconductor interface (Fig. 2-10b), and resultant linear i - V behavior.

In addition, metal-semiconductor contacts can also display nonlinear or rectifying characteristics, resulting in unexpected changes in circuit behavior. Schottky-barrier or metal-semiconductor diodes are based on the rectifying properties of certain metal contacts to semiconductors. If a band diagram is drawn for a metal with a work function larger than that of the n-type semiconductor it contacts; the resultant structure of Fig. 2-11 behaves very much like the p-n junction diode of Fig. 2-8 upon biasing. There is, of course, no band bending in the metal. A barrier potential of height Φ_B now governs the asym-

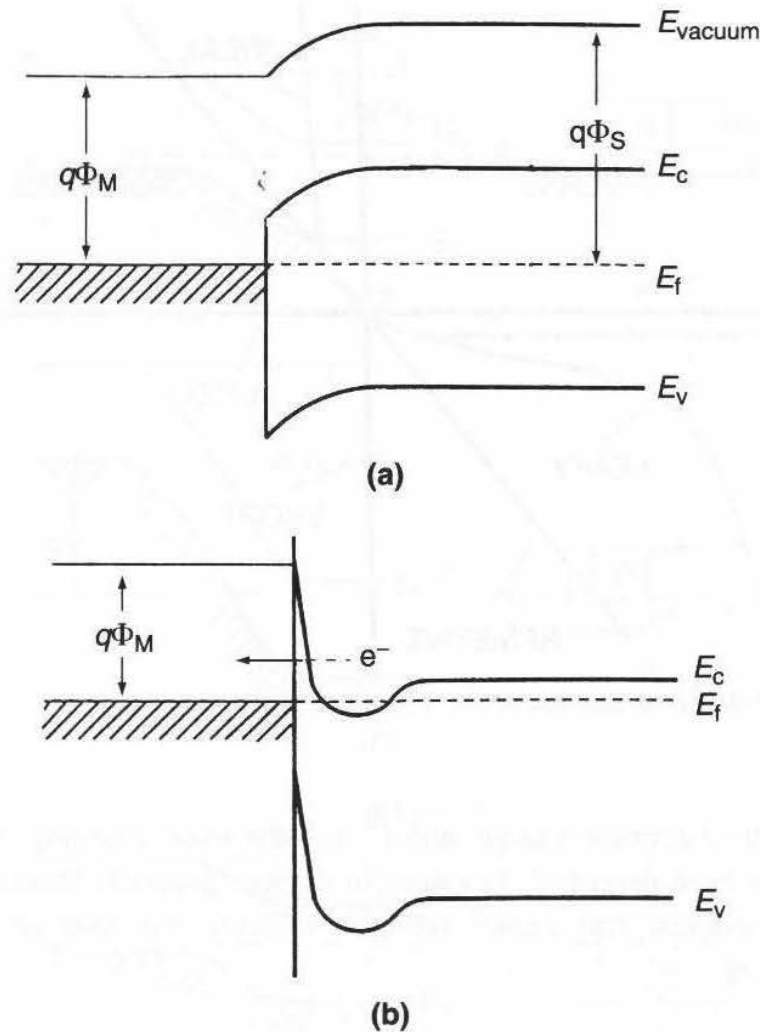


Fig. 2-10 (a) Band diagram representation for an idealized ohmic contact between a metal and semiconductor. (b) Band diagram for an ohmic contact illustrating tunneling. Φ_M and Φ_S are respective work functions.

metric transport of charge across the junction. Despite the fact that the i - V characteristics appear identical, Schottky-barrier and p-n diodes differ in one significant way. Under forward bias, majority electrons injected over the barrier are hotter than normal electrons in the metal. Their excess energy is very rapidly quenched, and such thermalized electrons cannot flow back into the semiconductor upon reverse bias because they face a large energy barrier in that direction. The short recovery time of 10^{-11} seconds makes Schottky diodes capable of very fast switching in microwave circuits.

Applications normally require either ohmic or Schottky behavior; interestingly, certain compound-semiconductor field-effect transistors contain both

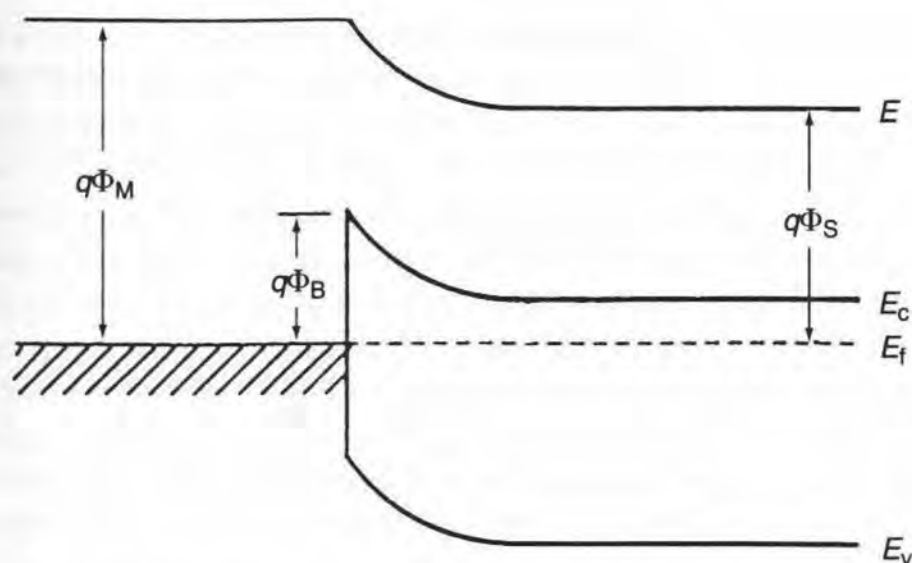


Fig. 2-11 Band diagram for idealized Schottky-barrier contact between a metal and semiconductor.

ohmic and Schottky contacts (Sect. 2.7.4). Irrespective of whether they are ohmic or rectifying, solid-state reactions between Al contacts and silicon as well as assorted metals (e.g., Au, Ti) and GaAs have been troublesome sources of degradation and failure in devices. These reliability problems are exacerbated when high current densities flow through contacts.

2.3.3 Deviations from Ideal Junction Behavior

There are a number ways that actual junction behavior differs from the ideal.

2.3.3.1 Characteristics. Rather than by Eq. 2-4, diode i - V characteristics are more nearly described by $i = i_R [\exp(\frac{qV}{n kT}) - 1]$, where n , the ideality factor, varies between 1 and 2. In addition, the contact potential due to work function differences between the involved semiconductors makes the i - V forward characteristics appear to be more squarish. Thus, very little forward current flows until a critical voltage somewhat below E_g is applied, after which the current rises very steeply.

Several effects serve to weaken rectifying behavior, resulting in the deviations from good i - V characteristics displayed in Fig. 2-9. Ohmic losses in diodes are manifested by passage of less current at high forward bias levels. Carrier generation outside of the junction region means higher reverse currents. These effects cause the diode to be more resistive and display linear i - V behavior as a limiting

case. In poorly made diodes the reverse current may sometimes have a surface component that leaks around the junction. Charged contaminants and semiconductor surface states foster electrical leakage paths. The proportions of bulk and surface currents depend on the electrical character of the junction at the point where the bulk meets the surface. Passivating Si diodes with SiO_2 layers minimizes such leakage. In good diodes the reverse current remains very small and then sharply increases at the point of breakdown, a subject discussed in the next section. Such characteristics are “hard,” distinguishing them from the “softer” i - V behavior exhibited by junctions possessing weak spots where localized breakdowns cause excessive leakage current.

Finally, there are the characteristics of the open diode that passes no current irrespective of voltage, and the shorted diode through which large currents flow at zero voltage.

2.3.3.2 Breakdown Diode breakdown is much less alarming than it sounds and does not necessarily mean damage. When the applied *reverse* voltage is increased beyond certain limits (Fig. 2-9) diodes undergo breakdown. Depending on doping level, electric fields in the junction ranging from 200 to 1000 kV/cm can develop in Si. Under such conditions minority carriers that produce the usual reverse or leakage current are accelerated to kinetic energies high enough for them to ionize matrix atoms upon collisional impact. In the case of a single incident electron, two electrons (the original one and the impacted valence electron) will now reside in the conduction band after collision. Each can impact-ionize lattice atoms repeatedly, and with sufficient electron multiplication, *avalanche breakdown* may occur. The reverse current rises to a value limited only by the external resistance of the circuit. In this state it may operate satisfactorily as a reverse-biased diode. However, if the reverse current is excessive, the junction can overheat, causing the diode to burn out. Zener diodes capitalize on operation in the breakdown state and are used as voltage references in rectifier and control circuits.

2.3.3.3 Punch-through Although the term sounds as ominous as breakdown, punch-through is not necessarily threatening to continued diode function. Punch-through occurs when lightly doped regions are also physically thin. Upon reverse biasing, the depletion region widens and may extend directly to the ohmic contact, thus shorting out the diode or transistor. Provided that the current flow is limited by the circuit resistance and no overheating occurs, the diode will recover its former properties when the reverse voltage is reduced. As we shall see later, punch-through also occurs in transistors.

2.4 BIPOLAR TRANSISTORS

2.4.1 Transistors in General

Transistors are the most important solid state devices. Because they are also the most numerous, device reliability and failure issues tend to be overwhelmingly focused on them. Since transistors contain two p-n junctions, have three contacts or terminals, are complex geometrically, and frequently require the action of perpendicular electrical fields, their behavior is more difficult to understand than that of diodes. To complicate matters further there are two fundamentally different types; in the bipolar junction transistor (BJT) current is amplified, while in the field effect transistor (FET) a voltage controls a current. The former rely on the participation of both types of carriers, hence the descriptor "bipolar." In contrast, only one carrier plays a prominent role in FETs. Transistors perform the basic electrical functions of amplification and switching in both discrete and IC configurations involving low as well as large power applications.

With regard to function it is helpful to distinguish between *linear* and on-off *digital* logic circuits. Linear circuits typically operate with low input voltage levels of perhaps $\sim 1\mu\text{V}$ or so (e.g., audio equipment) and outputs having small deviations about a quiescent operating dc level. Digital circuits are constrained to operate at fixed voltage levels (e.g., 0 or 5 V) and are impervious to noise levels that would present difficulty in linear circuitry. The latter are much less tolerant of variations in electrical component characteristics than is the case for digital circuits. With change in operating temperature, linear circuits may drift in and out of specifications, whereas digital circuits will either function properly or abruptly fail. Both BJTs and FETs are used in linear and digital circuits; computer logic and memory circuits are primarily composed of FETs, however, and these therefore dominate the transistor market.

2.4.2 Bipolar Junction Transistors

2.4.2.1 Operation It is common to start discussions of BJTs with the p-n-p type shown in Fig. 2-12 because the direction of majority carrier (hole) motion coincides with that of the current flow. There is also an n-p-n version of this device that is easier to fabricate. Since the electron mobility in Si ($\mu_e = 1450 \text{ cm}^2/\text{V}\cdot\text{s}$) is some three times higher than for holes, these faster n-p-n devices are preferred in most applications. By reversing the doping type, bias voltages, and current direction, the treatment applies to n-p-n devices as well. These three-terminal devices contain two junctions separating three semiconducting regions, i.e., the emitter, base, and collector. In the case under discussion these are doped p, n, and p, respec-

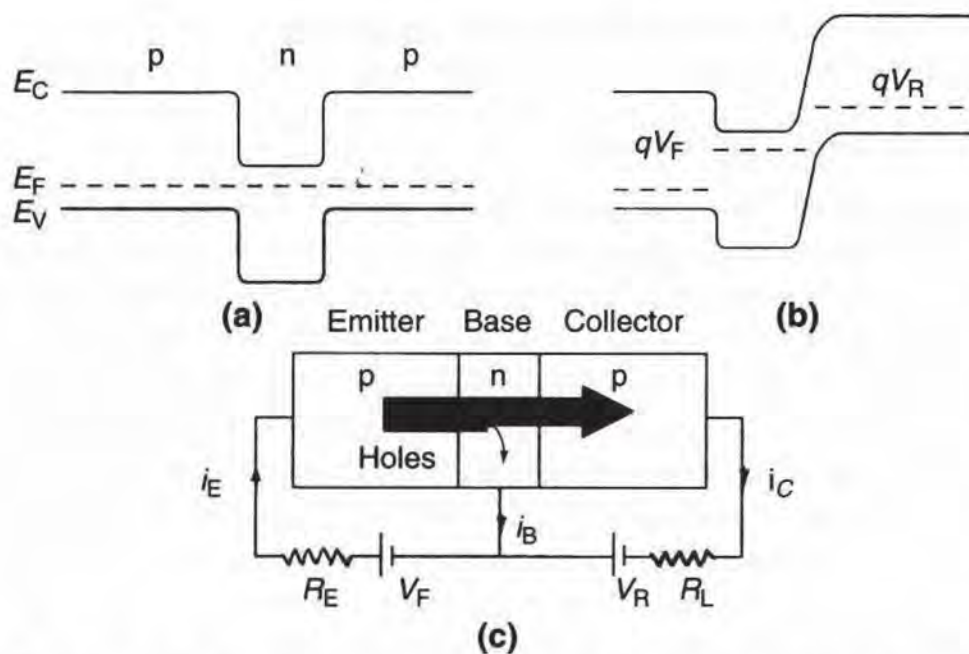


Fig. 2-12 (a) Band diagram representation of a p-n-p bipolar junction transistor in the common base configuration. Under zero bias, the Fermi levels line up across both junctions. (b) Band diagram representation of transistor under forward bias. (c) Transistor with bias voltages V_F and V_R applied.

tively, and are wired together in a common base configuration. When unbiased, the Fermi levels line up across the band diagram representation of the device (Fig. 2-12a).

During transistor operation the p-n junction on the left is forward biased, while the one on the right is reverse biased. The bands bend and move up or down as shown (Fig. 2-12b). First the heavily doped emitter injects majority (hole) carriers into the lightly doped base region. There the holes become minority carriers and would normally recombine with majority electrons. However, because the base is lightly doped and made very thin (e.g., 1 μm thick), holes diffuse right through it before they have a chance to recombine. Those holes that traverse the base-collector interface are now efficiently swept to the collector contact; equivalently, holes float uphill on the band diagram, which is their way of reducing energy.

In well-designed transistors the collector current, i_C , is only slightly smaller than the emitter current, i_E ; for example, i_B is typically 1 percent of i_E or i_C . Therefore, the common base current gain, defined as $\alpha = i_C/i_E$, is slightly less than unity. It is not difficult to calculate the ac power gain in the circuit shown in Fig. 2-12c. Due to the exponential character of Eq. 2-4, a small change in emitter-base voltage (ΔV_{EB}) will cause a large change (Δi_E) in emitter current. The power supplied to the forward-biased emitter junction, whose resistance is

R_E , is $R_E(\Delta i_E)^2$. Likewise, the power delivered to the load resistor R_L is $R_L(\Delta i_C)^2$, where Δi_C is the change in collector current. Because Δi_E and Δi_C are approximately equal, the power gain (G) is simply

$$G = \frac{R_L(\Delta i_C)^2}{R_E(\Delta i_E)^2} \approx \frac{R_L}{R_E}. \quad (2-5)$$

Typically, for $R_E = 25 \, \Omega$ and $R_L = 10 \, \text{k}\Omega$, a value of G equal to 400 can be anticipated.

Transistor action essentially occurs because the emitter-base voltage in one part of the device controls the collector current elsewhere in the device. In this way small changes in the current at one terminal enable the current through the other two terminals to be controlled. Similar transistor action occurs in other wiring configurations, e.g., common emitter. Specific doping levels and base dimensions are required for optimal transistor performance.

2.4.2.2 Planar Bipolar Transistor Discrete, or single bipolar, transistors of the type just discussed were the first to reach the market. Later the concept of integrating many bipolar transistors on a single wafer meant a new processing philosophy. Rather than stacking bulk semiconductor regions "side to side" as in Fig. 2-12, a "top to bottom" thin film technology was adopted. By sequentially doping through the wafer thickness, the n-p-n (or p-n-p) junction depths, and the base thickness in particular, can be very well controlled. The problem is how to make contact to the emitter, base, and collector regions. Obviously, contacts can only be made on the top surface of the wafer. Novel methods have been developed to direct the current flow, as shown in Fig. 2-13. Electrons injected from the emitter cross the narrow base region vertically down into a highly conducting, heavily doped buried layer that serves as a horizontal extension of the collector. They flow laterally through the buried layer and then veer up to the collector contact at the top surface. Provision must also be made to isolate (insulate) the transistor from other nearby circuit elements and devices. The operating principle of the planar bipolar transistor is the same as that for the discrete bipolar transistor. High operational speeds and current capacities have earmarked these devices for logic as well as input/output circuits.

The most complex digital system can be constructed from circuits that perform relatively few operations such as AND, OR, and INVERT. Individual circuits that perform these functions are known as logic gates. Over the years different logic gate circuits containing collections of bipolar transistors have been designed and incorporated into a wide variety of digital electronics. We just have time to mention some of the popular ones, namely emitter-coupled logic (ECL),

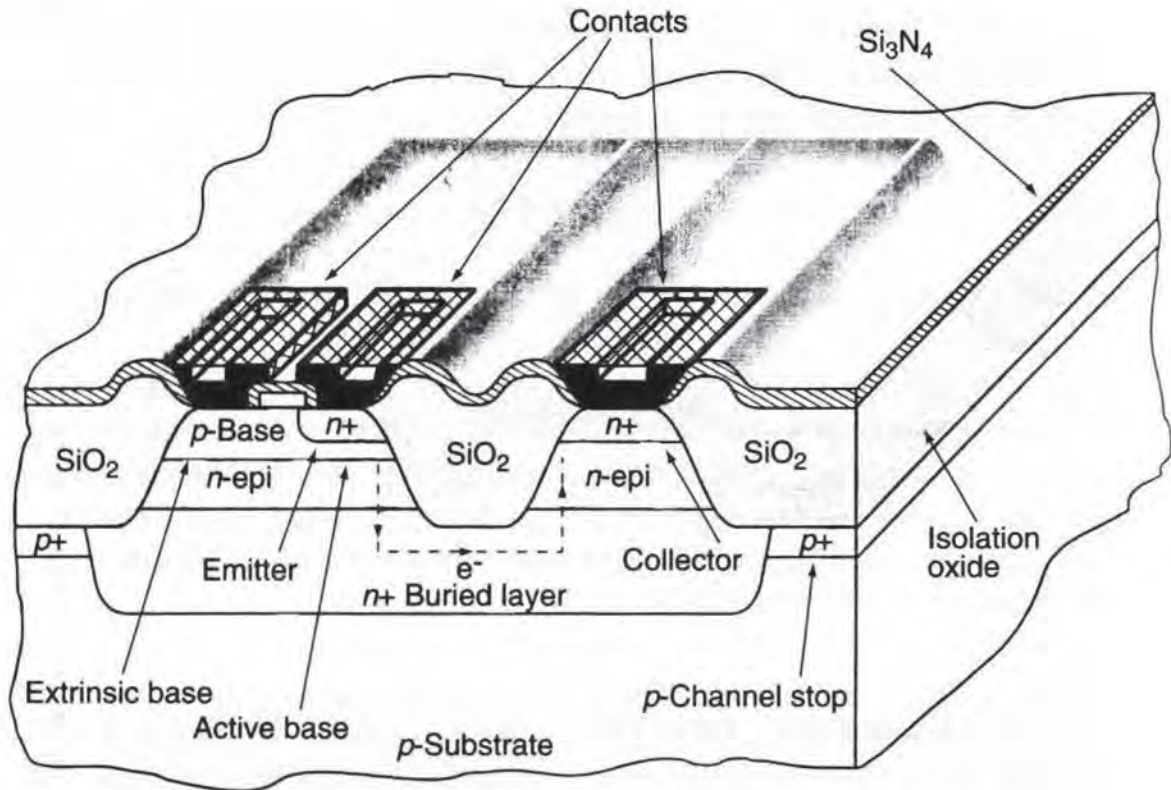


Fig. 2-13 Schematic of the planar n-p-n bipolar transistor structure utilizing oxide isolation. The current path from emitter to base to collector is indicated. From S. Middleman and A.K. Hochberg, *Process Engineering Analysis in Semiconductor Device Fabrication*, McGraw Hill, New York (1993).

transistor-transistor logic (T^2L), and integrated-injection logic (I^2L). The reader is referred to electrical engineering texts for descriptions of their operation and relative merits.

Logic functions can also be carried out more efficiently by field effect transistors, and it is to these devices that we now turn our attention.

2.5 FIELD EFFECT TRANSISTORS

2.5.1 Introduction

Like BJTs, field effect transistors (FETs) are three terminal devices; unlike them, a gate voltage rather than a current controls the current at the other two terminals. Field effect transistors are easily the most widely used of all solid-state

devices. They currently dominate computer logic as well as memory integrated circuit chips and are projected to widen this dominance. At the outset, two distinct types of field effect transistor must be distinguished. The first and most important is the metal oxide semiconductor (MOS) type, where the metal (M) gate electrode is separated from the semiconductor (S) by a thin insulator film (i.e., SiO_2). Voltage applied to the gate controls the current flow between underlying semiconductor regions by either creating or eliminating a conducting channel. Junction field effect transistors (JFET) are the second type, and they function by applying a gate voltage to vary the depletion width of a reverse-biased p-n junction.

2.5.2 The MOS Capacitor

Knowing how the MOS capacitor structure responds to applied voltage is prerequisite to understanding MOS field effect transistor (MOSFET) behavior. This capacitor structure, which is the essential building block of silicon MOSFETs as well as charge coupled devices, actually consists of two capacitors in series. The first has a fixed value of capacitance determined by the thin SiO_2 film sandwiched between planar metal (or other conductor, e.g., polysilicon or silicide) and semiconductor electrodes. Underlying this is the capacitance associated with Si, which varies depending on the extent of charge depletion in the semiconductor. Now consider the "ideal" MOS structure for p-type Si, shown in Fig. 2-14a. With no gate voltage or bias applied, Fermi level equilibration leaves all bands horizontal, or flat; this condition arises because the work functions of the metal and semiconductor are the same, and there is no oxide charge. When a negative gate voltage ($-V_G$) is applied to the metal, the energy of all of its conduction electrons is raised uniformly by an amount qV relative to Si. Holes are attracted to the Si/ SiO_2 interface, making Si there appear to be even more p-type than originally. This causes band bending at the interface (Fig. 2-14b), which graphically describes a state known as *accumulation*. The high interfacial hole density effectively serves to make the Si a positively conducting capacitor plate. For high $-V_G$ values, the MOS capacitance (C_0), which depends inversely on SiO_2 thickness, is large and solely due to the oxide.

As the polarity of V_G is reversed, the overall capacitance decreases. The reason is that holes are repelled from the Si/ SiO_2 interface. In this *depletion* regime Si acts like a dielectric with a capacitance governed by the width of the depletion region (Fig. 2-14c). The two series capacitors have a lower overall capacitance than that of either the semiconductor or oxide alone. Now the bands bend the other way; the slope of the oxide conduction band edge or electric field reverses sign, and near the interface the semiconductor acquires an n-like char-

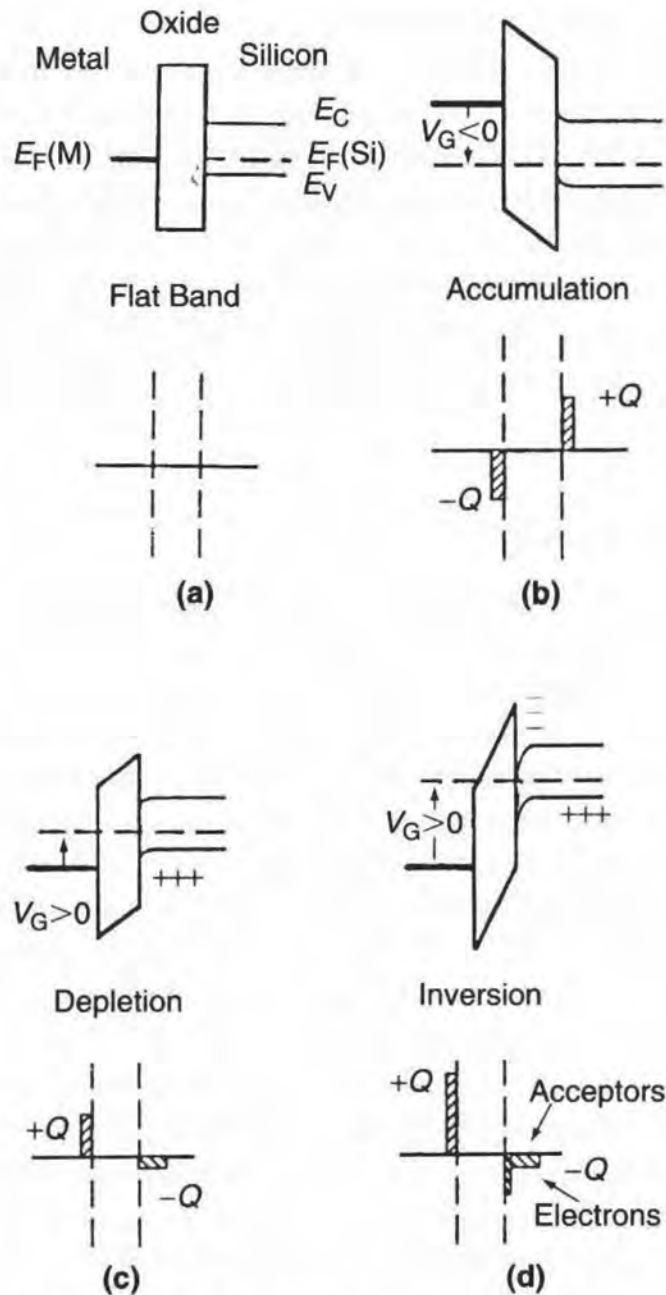


Fig. 2-14 (a) Ideal metal-oxide-p-type semiconductor (MOS) capacitor structure. Without applied gate voltage there is no oxide field, and a flat-band condition prevails. (b) Negative gate-voltage—accumulation of holes results in band bending at the interface. (c) Positive gate voltage—expulsion of holes causes Si to act like a dielectric with a capacitance governed by the width of the depletion region. (d) Large positive gate voltage—inversion regime is created and Si interface is effectively n-type.

acter. At still higher positive voltages, electrons are drawn to the interfacial region, where the semiconductor is locally inverted to n-type. Extreme band bending is characteristic of large values of V_G . An *inversion* regime (Fig. 2-14d) is created where the overall capacitance reaches a minimum. Effective charge distributions (Q) on the metal and in the semiconductor are depicted below each of the MOS bias states.

Ideal capacitance-voltage (C-V) characteristics derived from biasing the MOS structure are displayed in Fig. 2-15; an entirely complementary behavior is exhibited by n-type Si but with opposite gate voltage polarities. The MOS capacitor not only plays an active role in device function, but has proven valuable in assessing bulk ionic contamination (e.g., Na^+) and interfacial electronic charge (Q_i) introduced during gate-oxide processing. When such charge exists it can produce large voltage shifts (i.e., $V = Q_i/C_o$) in the C-V characteristics (Fig. 2-15). Charging of gate oxides during plasma processing or exposure to ionizing radiation produces

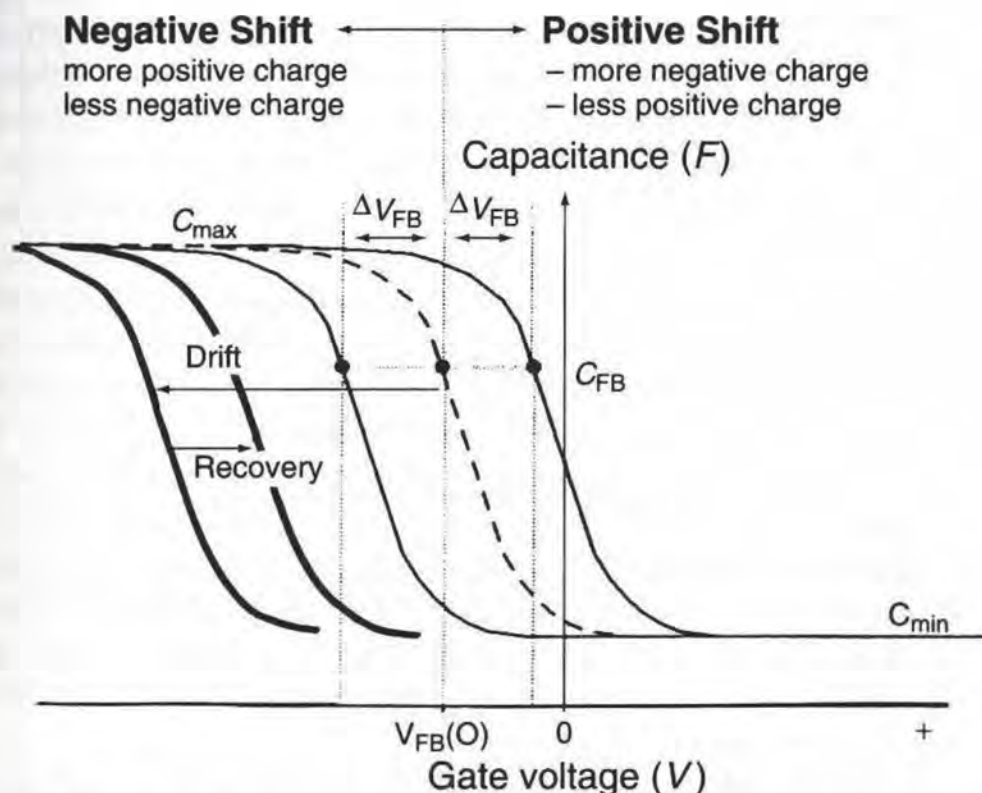


Fig. 2-15 Ideal capacitance-voltage (C-V) characteristics based on the biasing conditions in Fig. 2-14. Flat-band voltage (V_{FB}) shifts from uncharged capacitor (dotted) due to oxide charge are shown. More positive charge (less negative charge) produces negative voltage shifts, and more negative charge (less positive charge) produces positive voltage shifts.

similar voltage shifts. Charge drift and recovery in electric fields cause operational instabilities in devices.

2.5.3 MOS Field Effect Transistor

The n-type MOS (NMOS) field effect transistor is shown in Fig. 2-16a, where it is evident that n-doped source and drain regions flank a central MOS capacitor section composed of p-type silicon. Beneath the oxide there are essentially two p-n junctions connected back to back. When no gate voltage is applied, only a small reverse leakage current flows between the source and drain. Recalling the previous section, we note that application of a sufficiently large $+V_G$ causes inversion of the MOS structure. So-called “strong inversion” essentially occurs at a threshold gate-voltage level V_T given by

$$V_T = \Delta\Phi_{MS} - Q_i/C_o + Q_d/C_o + 2\phi_F. \quad (2-6)$$

At this voltage a thin surface inversion layer or conducting n-channel forms joining the two heavily doped n^+ regions, and the transistor is turned on. This equation basically expresses that bands must first be flat, and the effective voltage contributions to V_T by the first two terms ensure this. The work function difference between the gate electrode and semiconductor is $\Delta\Phi_{MS}$, while the second term was introduced above. Next, the p-type Si must be depleted of hole charge, Q_d (term 3), and finally, strong inversion (term 4) must occur. The quantity ϕ_F is essentially the difference in magnitude between the midgap (intrinsic) level and the actual Fermi energy, E_F . Both Q_d and $2\phi_F$ as well as E_F depend on doping level.

As written, this formula refers to n-channel devices, but it also applies to p-channel devices as well if the signs of the last two terms on the right are changed from $+$ to $-$. It should be noted that large contributions to V_T are made by both Q_i and oxide capacitance, C_o . Advantage is taken of this by controllably adjusting V_T through ion implantation in order to effectively change Q_i , and by growing thick oxides to electrically isolate devices. By the same token, even slight time-dependent electron charging of the oxide in service increases Q_i and hence alters V_T . In fact, the same factors that cause shifts in C-V characteristics induce threshold voltage shifts. These cause reliability problems because MOSFET action may be unintentionally triggered.

Once the underlying Si is inverted, the electron current flow from the source to drain through the surface n-channel can be controlled by varying the drain voltage (V_D). The remaining portions of Fig. 2-16 reveal what happens with successively higher drain voltages. If V_D is small, electron flow yields a drain current, $i_D = V_D/R$, where R is the channel resistance. This linear region is indicated on the $i_D - V_D$ characteristics of Fig. 2-17a. As V_D increases, it eventually

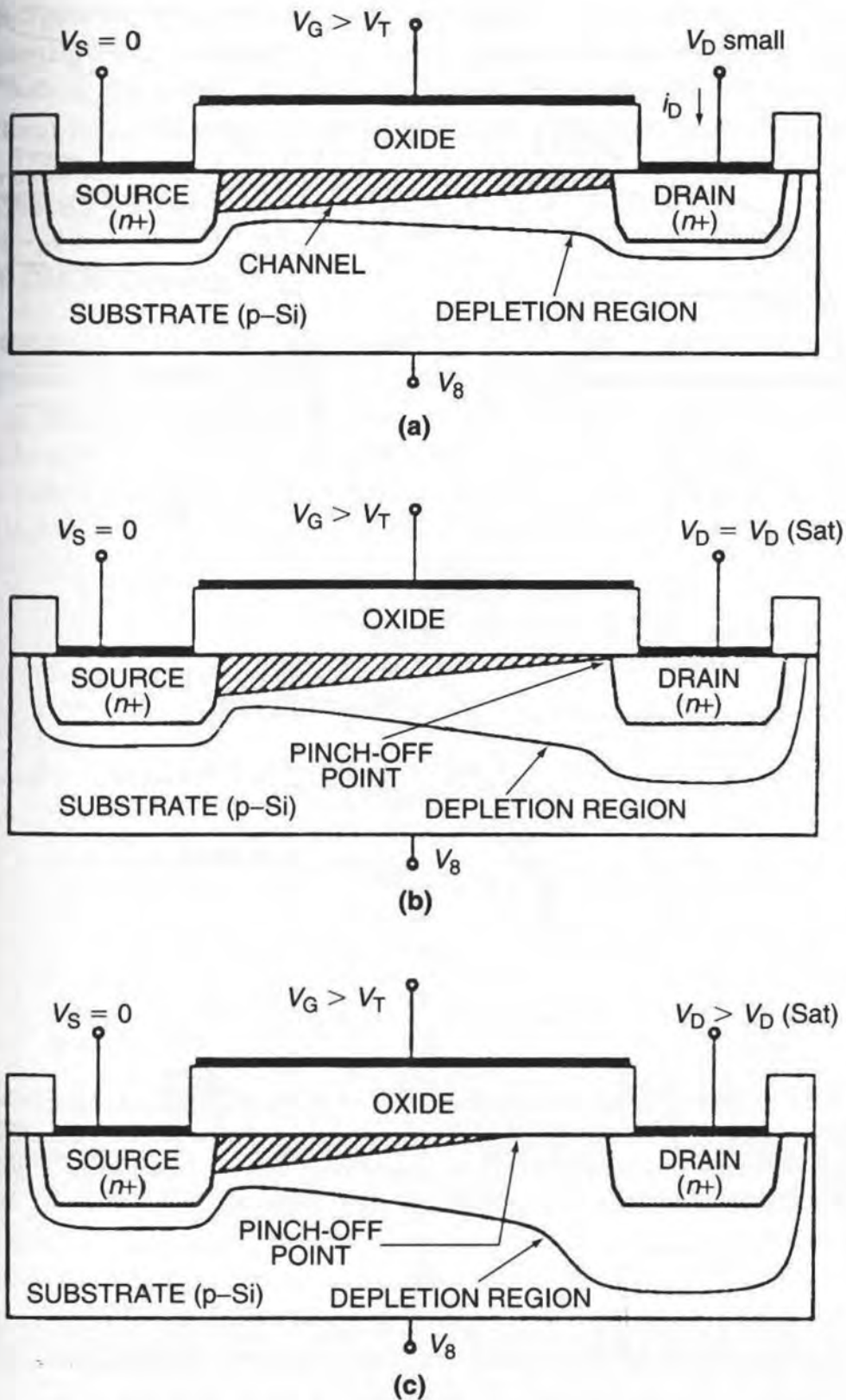


Fig. 2-16 Behavior of n-MOS field effect transistor. (a) Cross-sectional view showing operation in the linear region. (b) Operation at the edge of saturation (pinch-off). (c) Operation beyond saturation.

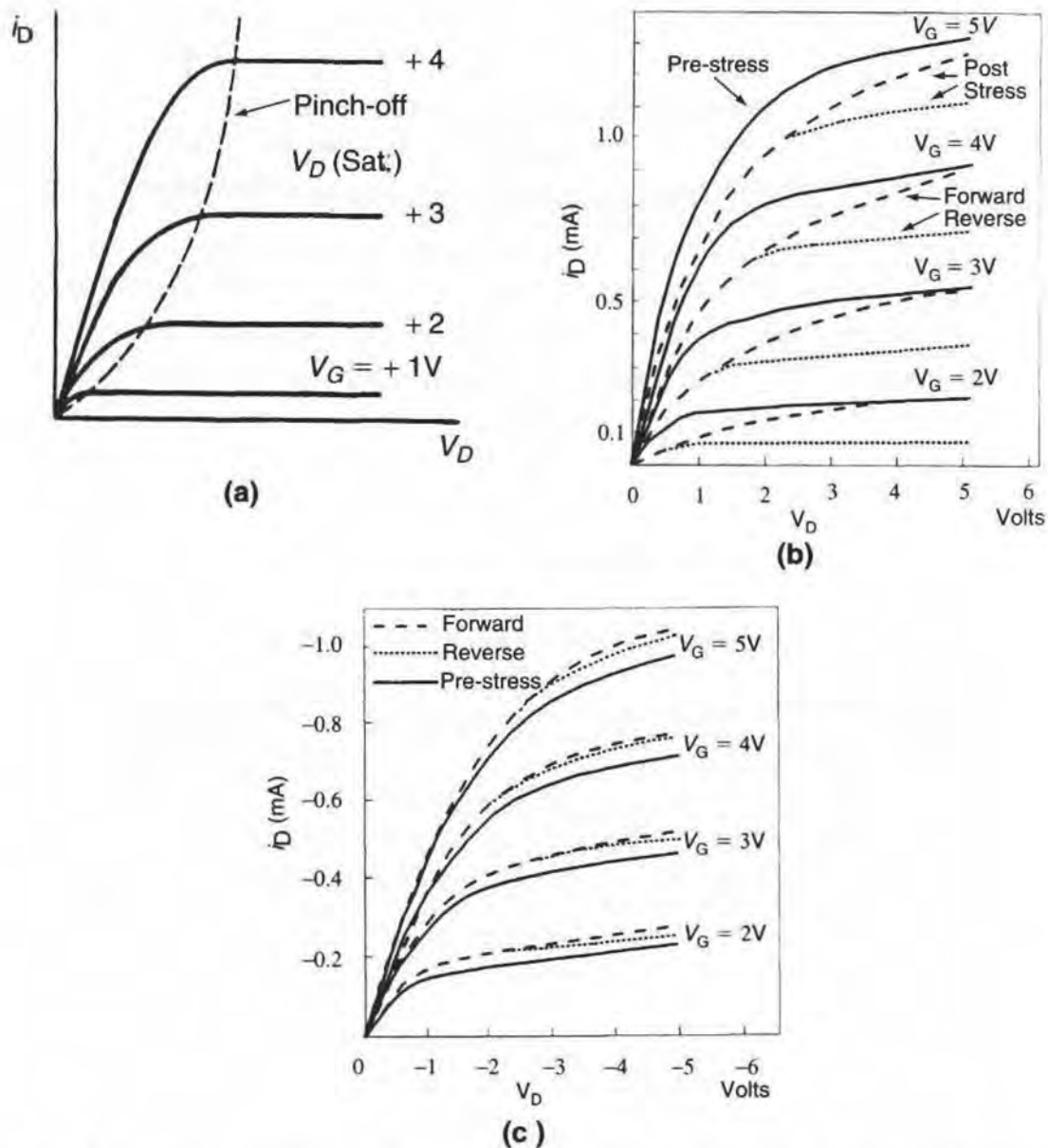


Fig. 2-17 Current-voltage (i_D - V_D) characteristics of the MOSFET. (a) Ideal behavior. (b) Forward and reverse n-MOSFET i_D - V_D characteristics after hot-carrier oxide damage. (c) Forward and reverse p-MOSFET i_D - V_D characteristics after hot-carrier oxide damage. From A. Schwerin, W. Hansch, and W. Weber, *IEEE Trans. on Electron Devices*, ED-34, 2493 (1987).

reaches a point where the inversion layer width at the drain shrinks to zero, the pinch-off state. Beyond pinch-off, i_D essentially saturates, but the channel width is effectively reduced. Electron injection into the drain depletion region is now similar to that of carrier injection from an emitter-base junction into the base-collector depletion region of a bipolar transistor. In addition to these n- and p-channel “enhancement mode” MOSFETs, there are “depletion mode” MOS-

FETS. These are fabricated containing the channel in the conducting state, and are thus switched on even when $V_G = 0$; during operation they must be turned off.

In both n- and p-type MOSFETs degraded by hot-carrier induced oxide damage, the i_D - V_D characteristics deviate from the prestressed state as shown in Figs. 2-17b and c. In essence, drain currents decline in n-MOSFETs and rise in p-MOSFETs. Hot-carrier effects are addressed again in Section 6-4.

2.5.4 CMOS Devices

The combination of MOSFETs, one n-type and the other p-type, is known as a complementary MOS or CMOS device (Fig. 2-18). By connecting gates in common as the input and drains in common as the output as shown, this device operates as an inverter in digital circuits. We first recall that V_T is positive in the n-channel transistor and negative in the p-channel transistor. When the input voltage (V_{in}) is "high," the n-channel device is in the "on" state and the p-channel

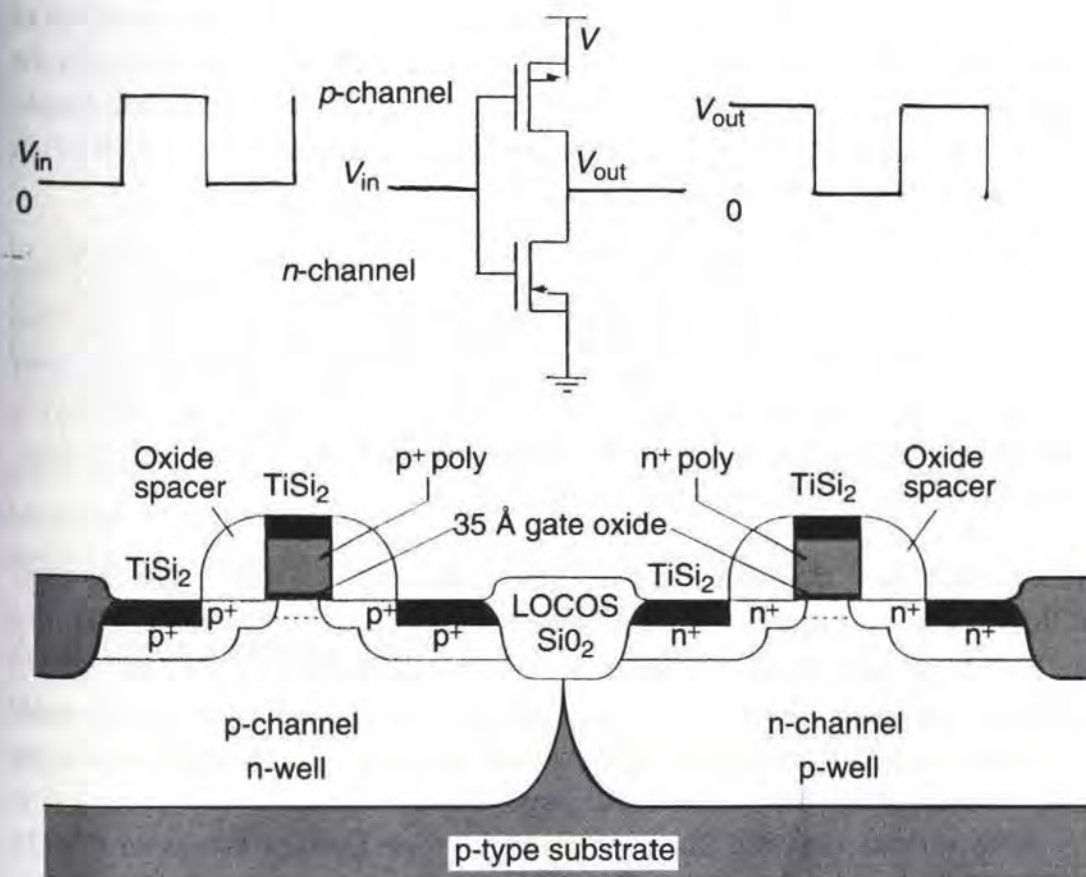


Fig. 2-18 **Top:** Circuit connection of transistors in a complementary MOS (CMOS) device. The indicated inverter function requires common gates as the input and common drains as the output. **Bottom:** CMOS structure.

device is in the “off” state. Therefore, the channel resistance of the p-device is much higher than that of the n-device. This causes voltage V to drop mostly across the p-channel, and the output voltage (V_{out}) is “low.” If the input voltage is now “low,” the n-device is off and most of V_{out} stems from V which represents the “high” state. Therefore, in either state a transistor is off, and there is no dc current path to ground other than leakage current. No drain current flows except during the very short switching period. The fact that little power is drawn is a primary reason why CMOS dominates the market for low- power logic and memory ICs where switching is the principal function.

2.5.5 MOSFET Instabilities and Malfunction

2.5.5.1 Oxide Charge The critical gate oxide is quite hospitable to both ionic and electronic charge. In addition to *mobile* alkali ion contaminants, there is both bulk oxide positive and negative *trapped* charge, as well as *interfacial* charge trapped at the Si-SiO₂ boundary, and *fixed* charge some 3 nm away. Area densities for the latter two are 10^{10} to $10^{11}/\text{cm}^2$. Sources of the charge are thought to be structural defects related to bond breaking and metallic impurity contamination in the oxidation process. During device operation, energetic, or hot, electrons in the channel may also get scattered into the oxide. Time-dependent dielectric breakdown and shifts in V_T are two important reliability consequences of such oxide charge buildup. This subject will be addressed more fully in Chapter 6.

2.5.5.2 Short Channel Effects A concern in VLSI and ULSI circuits is the existence of anomalous MOS transistors with shorter source-drain spacing than minimum design rules specify. The creation of a completely depleted area under the gate will then occur at lower gate voltages, and V_T will be effectively reduced. Electric fields are also higher in the drain region, making such transistors much more susceptible to hot electron effects.

2.5.5.3 Parasitic Bipolar Transistors: Latchup Despite efforts to confine MOS transistor action to a designated region of the substrate, sometimes unintended parasitic bipolar transistor behavior intrudes into the same or a neighboring location. So long as no p-n junction is forward biased, the bipolar transistor is inoperative. But for high enough substrate-source voltages, a large emitter current is injected, resulting in large device currents.

A more serious parasitic phenomenon known as *latchup* occurs in CMOS devices. Here the closely spaced sources and drains of n and p channel MOSFETs give rise to two parasitic bipolar transistors, one n-p-n and one p-n-p, as shown in Fig. 2-19. These transistors are cross connected so that the base-collector junctions are common. Under bias, the p-n-p collector delivers current

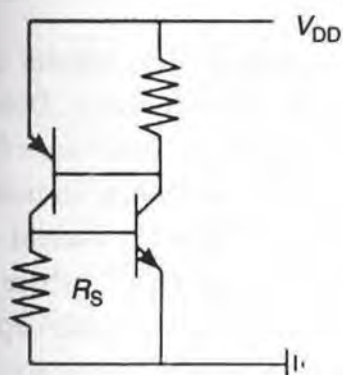
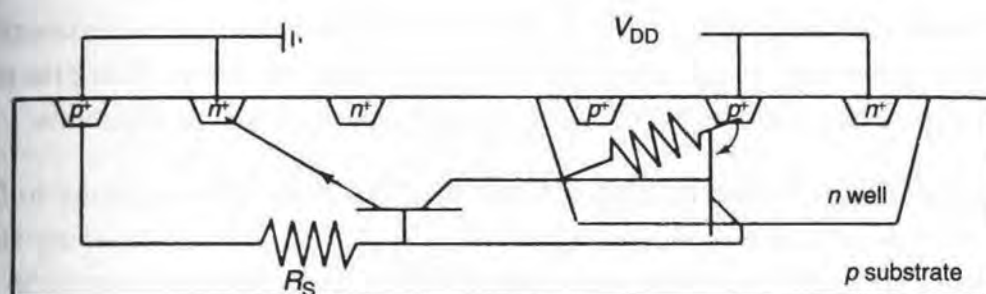


Fig. 2-19 Latchup arising in CMOS devices due the creation of closely spaced and coupled n-p-n and p-n-p parasitic bipolar transistors.

to the n-p-n base, and the n-p-n collector delivers power to the p-n-p base. If the gains of these transistors are high, both are driven into saturation, and the power supply voltages become connected through a low resistance path. In essence, a device known as a thyristor, which contains p-n-p-n doped regions, is effectively created. Carriers generated in unwanted regions by leakage currents, substrate current, or even photon irradiation can trigger latchup. When this happens, short-circuit currents rapidly overheat the IC chip, destroying it. Better design of transistor isolation is effective in combating latchup.

2.5.6 Bipolar versus CMOS

Now that the two major transistor technologies have been introduced, it is pedagogically worthwhile to compare the characteristics and attributes of each type of device (11).

a. Size. Using the same lithography tools, the footprint of bipolar transistors is typically 20–40 percent larger than the corresponding MOS transistor. A key advantage of CMOS technology is the smaller device area, which translates to a higher packing density.

b. Power Consumption. As indicated above, standby power consumption of CMOS transistors is much less than that for bipolar transistors. This plus high packing densities makes CMOS the obvious choice for ULSI applications.

c. Complexity of Fabrication Process. CMOS devices are easier to fabricate than bipolar transistors with typically 20 percent fewer processing steps required. Part of the simplicity and smaller size of MOS devices stems from the less stringent isolation required. To isolate or insulate devices from one another on the same chip, more complex and deeper dielectric structures (trenches) are required for bipolar than for MOS devices.

d. Current Carrying Capacity. An important advantage of the bipolar transistor is its decidedly superior gain because of its high transconductance. Defined as the derivative of the output current to input voltage, the transconductance (G_m) of a transistor is a measure of its quality as an amplifying device. In bipolar transistors, $G_m = q i_c / (kT)$, where q is the electronic charge, i_c is the collector current, and kT has the usual meaning. For MOS transistors $G_m = (2\mu i_{DS} C_{ox} Z/L)^{1/2}$, where μ is the electron mobility, i_{DS} is the drain-source current, C_{ox} is the gate-oxide capacitance, and Z and L are the channel length and width, respectively. Substitution of typical values reveals that to match the bipolar transconductance, the dimensions of MOS transistors must be a few orders of magnitude larger.

e. Sensitivity and Speed. The exponential dependence of the current on the base-emitter voltage gives the bipolar transistor superior signal detection capabilities relative to MOS devices. Bipolar devices are also faster than MOS transistors and exhibit higher frequency response and gain. This makes them well suited for high performance analog applications.

f. Market Share. Currently, NMOS and CMOS technologies account for about 75% and bipolar 20% of the IC market share. Compound semiconductors and BiCMOS make up the remainder.

For flexible circuit design, BICMOS technology combines the high density of CMOS with the speed of bipolar transistors. This technology has been used in the Intel Pentium microprocessor, a chip containing 3.3 million transistors. In general, BICMOS circuits employ CMOS transistors to perform the logic functions and bipolar transistors to drive the output loads.

2.5.7 Junction Field Effect Transistor (JFET)

Like the MOSFET, the JFET, depicted in Fig. 2-20a, also consists of a gate-source-drain combination. Rather than an oxide/semiconductor structure, the active region beneath the gate consists of three vertically stacked semicon-

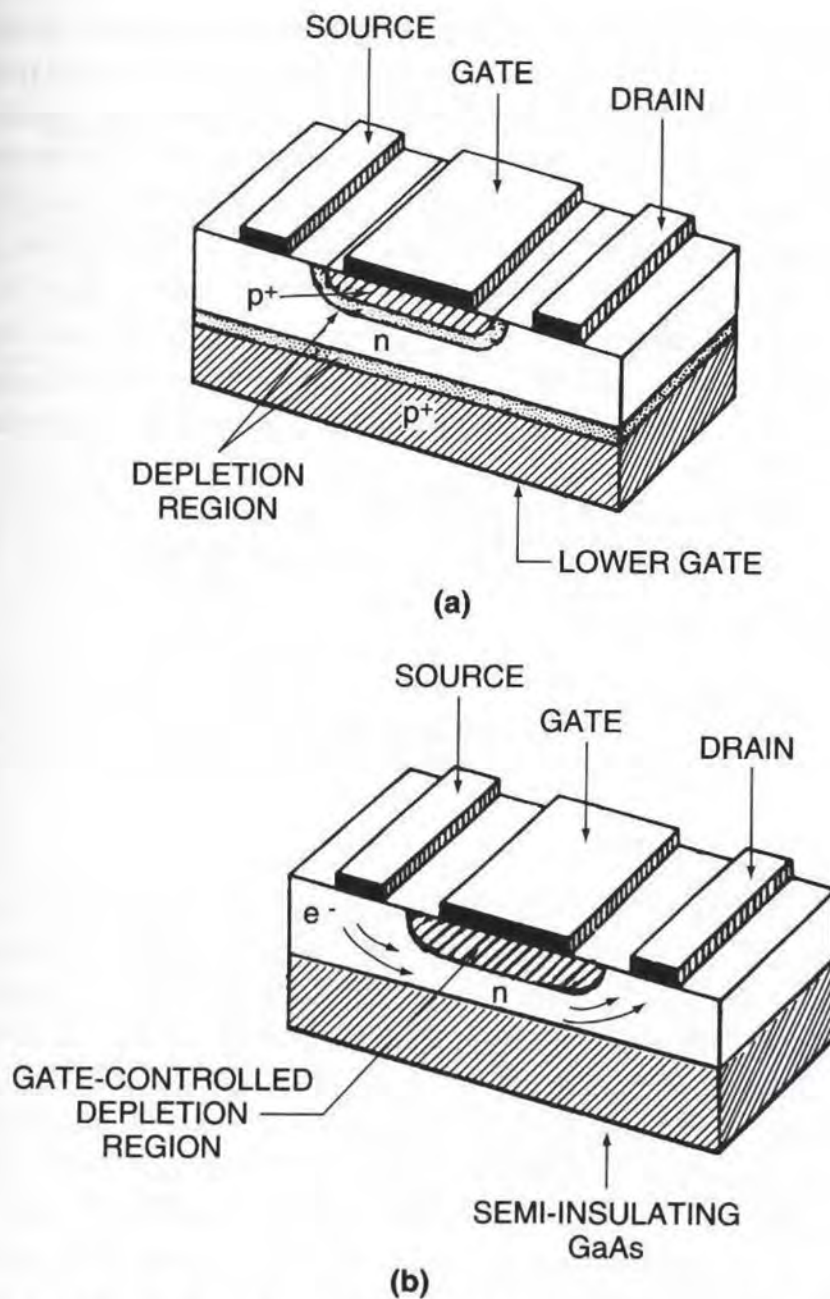


Fig. 2-20 (a) View of a junction field-effect transistor with a detail of the central channel region. From S.M. Sze, *Semiconductor Devices, Physics and Technology*, J. Wiley and Sons, New York (1985), Copyright by Bell Telephone Laboratories, Inc. (b) Schematic of the MESFET. Current flow through the channel is controlled by the gate voltage, which varies the depletion layer width.

ductor layers. In operation the upper and lower gates (the top and bottom heavily doped p-layers) are connected together, while the the lightly doped n-channel in between has the drain voltage (V_D) applied across it. Both p–n junctions are slightly reverse biased initially, so that depletion layers extend into the channel. As $+V_D$ is applied, a channel current (i_D) flows whose magnitude is simply given by Ohm's law, or V_D divided by the channel resistance. At larger values of V_D , the voltage drop (i.e., bias) that progressively rises from source to drain increases both i_D and the depletion width in concert. Finally, the two depletion regions merge at the drain, and the channel pinches. At that point the current saturates just as it does in MOSFETs and the i_D – V_D characteristics resemble those displayed in Fig. 2-17a.

Similar devices known as MESFETs are widely used in high-frequency circuits. Fabricated in III–V compound semiconductors, they are discussed in Sect. 2.7.3.

2.6 MEMORIES

2.6.1 Types of Memories

Together with logic chips, memory chips are critical for computer circuit applications. Memories are circuits that enable the storage of digital information, or data, in terms of bits (binary digits). They now account for close to 40% of the integrated circuit market and can stand alone on a single IC chip or be embedded within a larger IC. A road map through the broad variety of memories used, and the categories they belong to, is shown in Fig. 2-21. Memories that lose their data when the supply voltage is removed are said to be *volatile*, while those that retain data under such conditions are *nonvolatile*. Another subdivision reveals that there are three kinds of memory—serial memory, random access memory (RAM), and read only memory (ROM). Volatile memories include serial, static (SRAM), and dynamic (DRAM) types. Three-quarters of the total memory market, including both volatile and nonvolatile, is dominated by DRAMs. In serial memories, data enters and leaves in the same sequence, so the “first in” is the “first out.” Charge coupled devices (CCD) described later are used in such memories for video applications.

Random access memories, which come in SRAM and DRAM varieties, are arranged in a matrix array and require both read–write data and address inputs and data outputs. Data can either be written or read directly in a random way, unlike memory locations on magnetic tape, which must be lined up with the read–write head. Storage of data occurs within memory cells. In SRAMs these cells contain

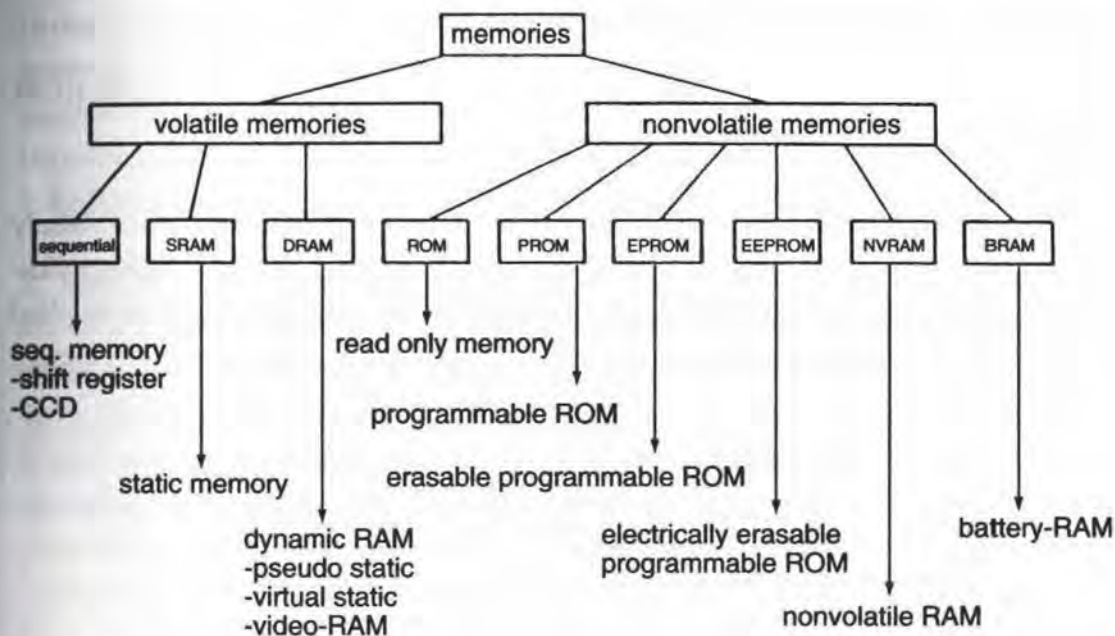


Fig. 2-21 Overview of the different types of semiconductor memories. From H.J.M. Veendrick, *MOS ICs-From Basics to ASICs*, VCH Publishers, Weinheim (1992).

inverter circuits, and combinations of them are known as flip-flops, which are capable of assuming two states. While SRAMs are the fastest of all MOS memories, this is accomplished at the expense of requiring four to six transistors per cell. In contrast, DRAM cells contain an integrated transistor–capacitor unit, enabling greater packing densities at lower cost. The operation of a DRAM will be described in the next section. Since charge is stored in a capacitor within DRAM cells, gradual leakage necessitates refreshing approximately every 2-16 milliseconds. In contrast, static RAMs ensure that voltage levels are maintained as long as the power supply is on.

Read only memories are, in fact, random access memories that are permanently written during the fabrication process. Therefore, the information is lasting (nonvolatile) and can be read but not altered. There are also several kinds of ROMs that can be programmed by the user.

1. PROMs are ROMs that can only be programmed once. They usually consist of bipolar transistors, and the act of programming causes fuse links to be electrically blown, permanently fixing the memory circuit.
2. EPROMs are similar to PROMs but are erasable by exposure to UV radiation. Their operation will be discussed in the next section.
3. EEPROMs are PROMs that are electrically erasable and can be reprogrammed repeatedly.

2.6.2 Memories Are Made of This!

In this section the structure and operation of typical serial, RAM, and ROM memory devices will be briefly described.

1. Charge Coupled Device (CCD). The CCD of Fig. 2-22 consists of closely spaced MOS capacitors arrayed on a continuous SiO_2 layer that covers the underlying p-type semiconductor. If sufficiently large positive voltage pulses are applied to all of the electrodes, each capacitor will be depleted. In order to transfer charge

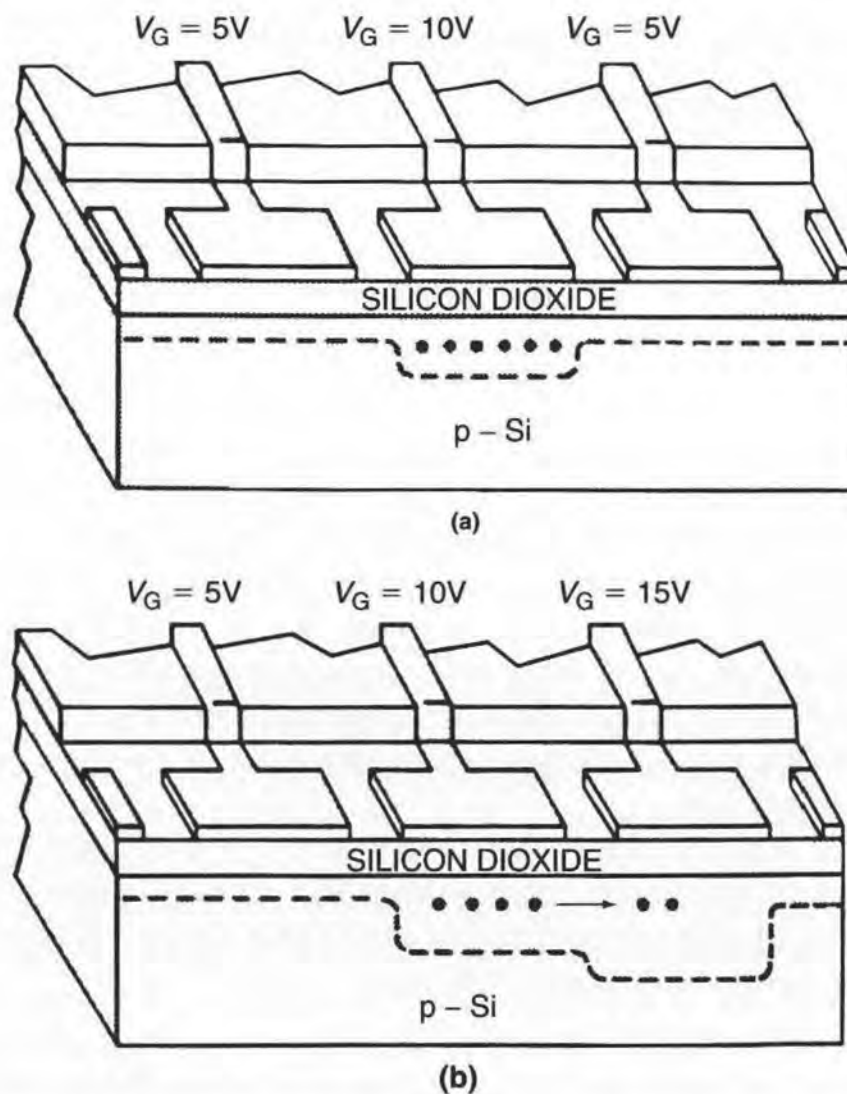


Fig. 2-22 Schematic of a charge-coupled device. Charge originally stored on center capacitor is transferred to neighboring cell by increasing bias voltage. From S.M. Sze, *Semiconductor Devices, Physics and Technology*, J. Wiley and Sons, New York (1985), Copyright by Bell Telephone Laboratories, Inc.

to one of the cells, a higher bias is applied there to create greater depletion. In effect, a potential well develops there that captures injected electrons. By pulsing the neighboring cell to the right more positively, electrons can be made to transfer between cells. This lively transfer of charge can proceed in both directions. CCDs have found use in imaging devices that ultimately produce electrical video signals.

2. DRAM. The drive to minimize the number of devices per memory cell has led to the widespread use of the NMOS transistor/capacitor combination. A cell cross-section and the way it is wired into the circuit are shown in Fig. 2-23. In 4Mb DRAMs the cell size is $<10 \mu\text{m}^2$, while in 16Mb DRAMs, areas less than $5 \mu\text{m}^2$ are involved. But, for 256Mb DRAMs the cell size is only $0.6 \mu\text{m}^2$! Shrinkage of DRAM cells mean that thicknesses of gate and capacitor oxide dielectrics must scale accordingly (12). The drain of the MOSFET serves as the conductive path between the inversion layers under the capacitor and the transfer gates. Electrodes of the storage capacitor are the drain/conducting polysilicon fill and the substrate ground; they sandwich the SiO_2 dielectric in between. High packing density requirements have motivated clever approaches to simultaneously reduce chip surface area while increasing charge storage capacity. By fabricating deep trench capacitors, the capacitance is markedly increased. Trench technology poses processing challenges in filling deep holes without the formation of defective cells. We shall return to the issue of defects in these important devices in Sect. 3.2.5.

3. EPROM. These read only devices are basically MOS transistors that contain an additional polysilicon gate layer embedded within the oxide beneath the gate electrode. Because the embedded gate is unconnected it is electrically floating. Through application of a high source-drain voltage pulse, an avalanche of electrons is initiated near the pinched-off region. Some electrons acquire enough energy this way to enter the oxide and charge the floating gate. It is for this reason that these MOSFETs are known as floating avalanche, or FAMOS, devices. In packaging these devices a window is installed over the chip to enable UV light to shine through and erase the memory. By locating the floating gate only a few nanometers from the silicon substrate, tunneling currents can charge the floating gate and also erase information; this is done in EEPROM devices.

2.6.3 Reliability Problems in Memories

Memory reliability problems can be divided into three categories (13), namely 1. processing or fabrication related, 2. operation related, and 3. environment related. Most concerns are common to SRAMs and DRAMs and to a lesser extent to EPROMs.

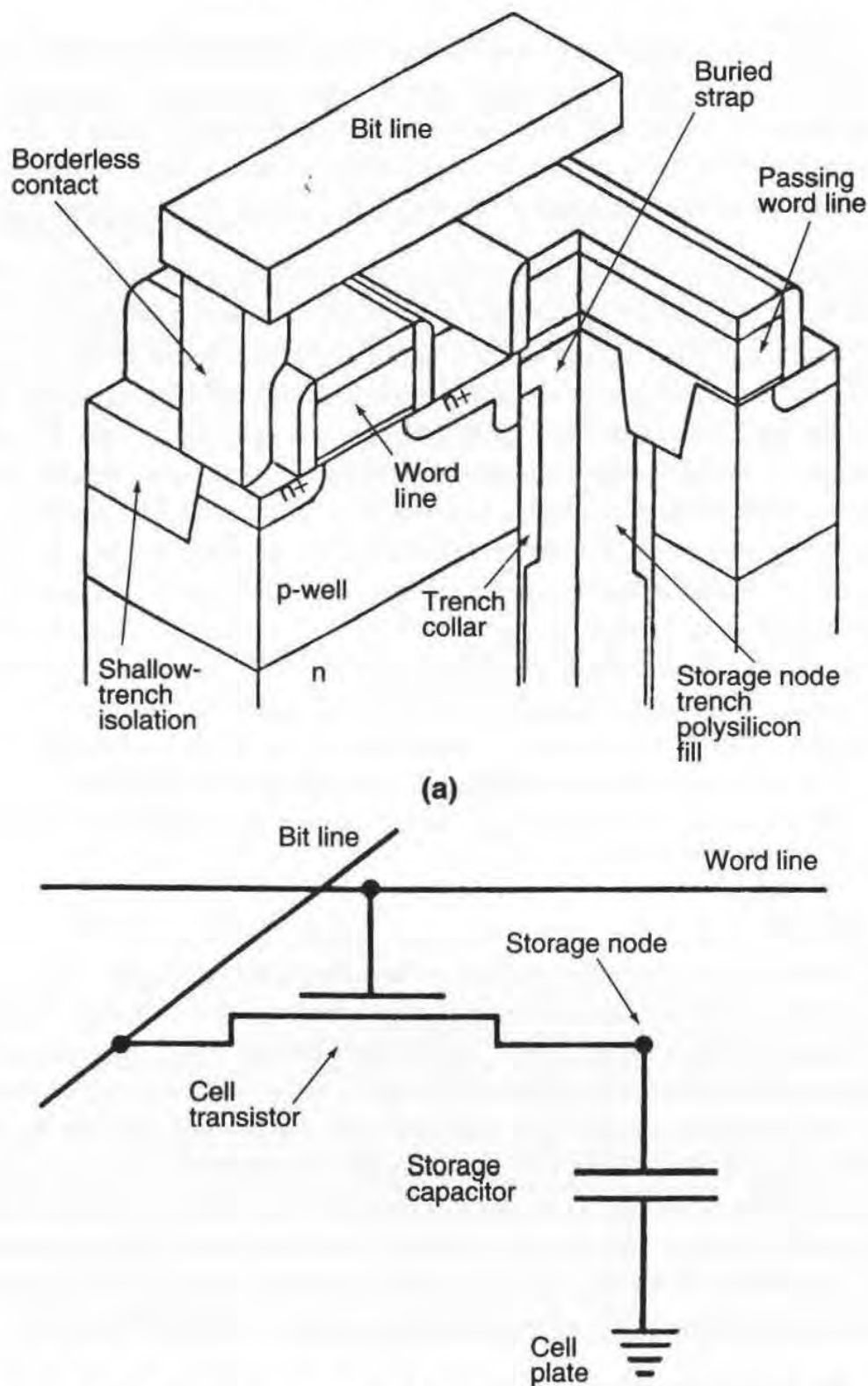


Fig. 2-23 Cross-section of a DRAM cell and its method of interconnection in the circuit. To write, the bit line is driven to a high or low logic level with the cell transistor on. Then the transistor is shut off leaving the capacitor charged high or low. (Charge leakage requires refreshing.) To read, the bit line is left floating as the transistor is turned on; the small change in bit-line potential is sensed and amplified to a full logic level. From reference 12.

Processing difficulties include control of the conductor linewidth and semiconductor doping levels, defects due to contamination, stress, and process-induced radiation. Yields are adversely impacted, and poor performance and degradation often show up in the infant mortality population. Adding redundant columns or rows of devices that can be programmed to substitute for failing memory cells is one of the strategies adopted to combat failed devices.

Operational reliability problems stem from the nonuniform scaling of device features to achieve high packing densities. Reduction in transistor gate lengths without corresponding changes in operating voltages has resulted in higher drain electric fields. These promote time-dependent changes in dielectric properties and threshold voltage shifts.

Environmental damage is largely related to packaging issues. On the chip level, nuclear radiation in the form of gamma rays and alpha particles, which may originate in the packaging materials (e.g., Pb-Sn solder), impinge on memory cells and ionize atoms, causing local soft errors. Chapter 7 addresses such reliability issues.

2.7 GAAS DEVICES

2.7.1 Why Compound Semiconductor Devices?

Hailed as the electronic materials of the future, advances in silicon technology have perpetually consigned compound semiconductors to a second-place status. Nevertheless, there are electronic functions that are performed inefficiently or not at all by Si devices; this is the reason for our interest in III-V materials, specifically GaAs, InP, and ternary and quaternary alloys based on these materials. A comparison with Si indicates several important reasons for their use in devices.

1. The speed (frequency) of devices made from compound semiconductors exceeds that for Si devices simply because electron mobilities are higher. For example, Table 2-2 reveals that $\mu_e(\text{GaAs})$ is some six times that for $\mu_e(\text{Si})$. This shortens electron transit times and enhances high-frequency operation. Very high speed microwave devices and integrated circuits require these attributes of GaAs devices.
2. The larger band gap in GaAs relative to Si means that it can withstand higher working temperatures, an attractive property in small geometry power

- devices used in microwave applications. Low power devices that operate at room temperature have correspondingly low thermal generation of carriers and leakage currents, hence low noise.
3. Direct band-gap transitions in compound semiconductors make optical emission and absorption processes more efficient than in Si (Sect. 2.2.2.4). This has enabled the production of efficient lasers, light emitting diodes, and photodetectors.
 4. Optical devices can often be made to operate at specific wavelengths within a broad spectral domain. In Si, only the wavelength corresponding to the band-gap energy is easily accessible. By alloying compound semiconductors, ternary and quaternary alloys emerge with energy gaps controllably tuned to a desired operating wavelength. Optical communications technology capitalizes on this to create high-performance light sources and detectors that operate at wavelengths where minimum optical losses occur.
 5. A desirable attribute of GaAs is its ability to be produced in semi-insulating form with resistivities of $\sim 10^8 \Omega\text{-cm}$. Such insulating single crystals make excellent substrates for GaAs microwave devices and integrated circuits.

Advantages seldom come without drawbacks. In GaAs we are dealing with a brittle, expensive material that is not easy to process and fabricate into devices. On the other hand, Si is mechanically robust and cheap, and significantly easier to process than GaAs. Silicon can also be protected and passivated by an easily grown, high-quality SiO_2 film, something that cannot be done in GaAs.

The subsequent discussion reflects the division of compound semiconductor devices according to microwave and opto-electronic applications. General tutorial reviews of the operation, reliability, and failure of compound semiconductor devices can be found in references 14, 15, and 16.

2.7.2 Microwave Applications

Microwave frequencies range between 10^9 Hz (1 GHz) to 1000 GHz with respective wavelengths of 30 to 0.03 cm. Within this spectral domain are a number of communication systems applications that are important in both the military and civilian sectors. Military applications include ground and airborne radar, electronic warfare including guided weapons, and satellite communications; civilian applications include microwave radar (for police, small boats, intruder alarms, and door openers), direct broadcast satellites (12 GHz) and mobile (1–3 GHz band) as well as cellular (~ 1 GHz) communications. To make

these systems function, devices that generate, detect, and amplify microwave signals are required. Since the limit for silicon devices is ~ 1.5 GHz, GaAs devices are preferred above this frequency. Microwave circuits are sometimes difficult to design because conductors such as stray solder act like antennas at these frequencies.

2.7.3 The GaAs MESFET

By replacing the upper p-n junction of a JFET (Fig. 2-20a) by a reverse-biased Schottky barrier, the MESFET emerges. The first part of the acronym for this device, which is deposited on a semi-insulating GaAs substrate, stands for (Schottky) metal (contact) epitaxial (compound) semiconductor. In the MESFET, schematically depicted in Fig. 2-20b, source-gate-drain electrodes combine to control current flow through a lightly n-doped single crystal (epitaxial) film of GaAs that serves as the channel. Sandwiching the channel are the metal contacts above and the semi-insulating GaAs substrate below. Device behavior and characteristics parallel those of the Si JFET. But GaAs MESFETs differ from their lower-frequency Si cousins because the channels are thinner ($\sim 0.1 \mu\text{m}$) and gate lengths shorter ($< 1 \mu\text{m}$). For channels of length L_c , the device cutoff frequency is given by $v_s/(2 \pi L_c)$, where v_s is the electron saturation velocity ($\sim 2 \times 10^7$ cm/s in GaAs). Thus for a $1 \mu\text{m}$ channel length, ~ 35 GHz operation is possible.

In addition to the low-power, small-signal MESFETs used in receiver front ends and digital integrated circuits, there are power MESFET devices that operate at higher voltages. These devices contain an interdigitated, or comb-like, array of contacts that encompasses a larger area with a deep active channel. The contrast in geometries and operating conditions is reflected in different failure mechanisms for low-signal relative to high-power devices.

In contrast to the Schottky gate contact, source and drain contacts in these transistors must be ohmic (e.g., Au-Ge). Accelerated high-temperature testing of these devices has revealed failures due to the degradation of the metal employed in these contacts. "Gate sinking," a common failure mechanism involving metal interdiffusion into the GaAs, is discussed further in Chapter 5. Decrease in drain current at zero gate bias, reduction in pinch-off voltage, and degradation of channel properties (e.g., increase in drain to source resistance) are some of the reliability problems associated with GaAs transistors. Improvements in gate metallizations over the past two decades have extended device lifetimes from thousands of hours employing Al to millions of hours with TiPtAu.

2.7.4 High Electron Mobility Transistor (HEMT)

These quantum devices operate much like MOSFETs and exhibit gains and speeds that are among the highest for any transistors (17, 18). A common figure of merit for FETs is the transconductance, and it can be increased by reducing the channel resistance through heavy doping. But this strategy degrades electron mobility because of carrier scattering by the ionized impurities. In HEMTs, the dual desiderata of high carrier concentrations and mobilities are simultaneously achieved. This is basically accomplished at the interface between n-doped AlGaAs and undoped GaAs, as indicated in Fig. 2-24a. Electrons in the conduction band of the doped AlGaAs (larger energy band gap) spill over into the GaAs (smaller energy band gap), where they are now trapped in a quantum-mechanical well that is free of ionized donor scatterers. As a result of band bending at the junction, the electrons in the well are essentially confined to a thin sheet. For this reason we speak of a two-dimensional electron gas, i.e., 2-DEG. An electric field applied normal to the page will result in enhanced carrier transport in the well channel. The effect of this so-called modulation doping is magnified at low temperatures because thermal scattering by vibrating lattice atoms is minimized. Incorporated within the structure of the GaAs-base HEMT shown in Fig. 2-24b is such a quantum well. Today, HEMTs also known as MODFETs have been fabricated with gate-lengths of approximately $0.25\text{ }\mu\text{m}$ in InP-base semiconductors. These devices demonstrably outperform GaAs MESFETs with respect to gain and low noise. Short gate-length GaAs and InGaAs MODFETs have frequency responses above 100 GHz with switching delays of ~ 5 picoseconds (15).

Sinking of Schottky gates and degradation of ohmic source and drain contacts are major failure mechanisms in both MESFETs and HEMTs. Electromigration, surface oxidation of GaAs, burnout, diffusion of dopants from channels, and intermetallic phase formation in Al–Au metallizations (14) are additional degradation mechanisms.

2.7.5 GaAs Integrated Circuits

Many of the same reasons for the integration of discreet Si devices have driven the development of both analog and digital memory and logic GaAs ICs. Except for microwave applications, these products have lagged behind their Si counterparts. In addition, there have been imperatives to fabricate monolithic microwave (and millimeter wave) integrated circuits, or MMICs. But microwave circuits are not simply large collections of FETs; in fact, some MMICs do not contain FETs at all! In some MMICs the number of passive circuit elements (resistors, inductors, diodes, transmission lines, interdigitated and overlay capacitors) may outnumber the FETs by a factor of ten to one (19). Recently, an extensive evaluation of

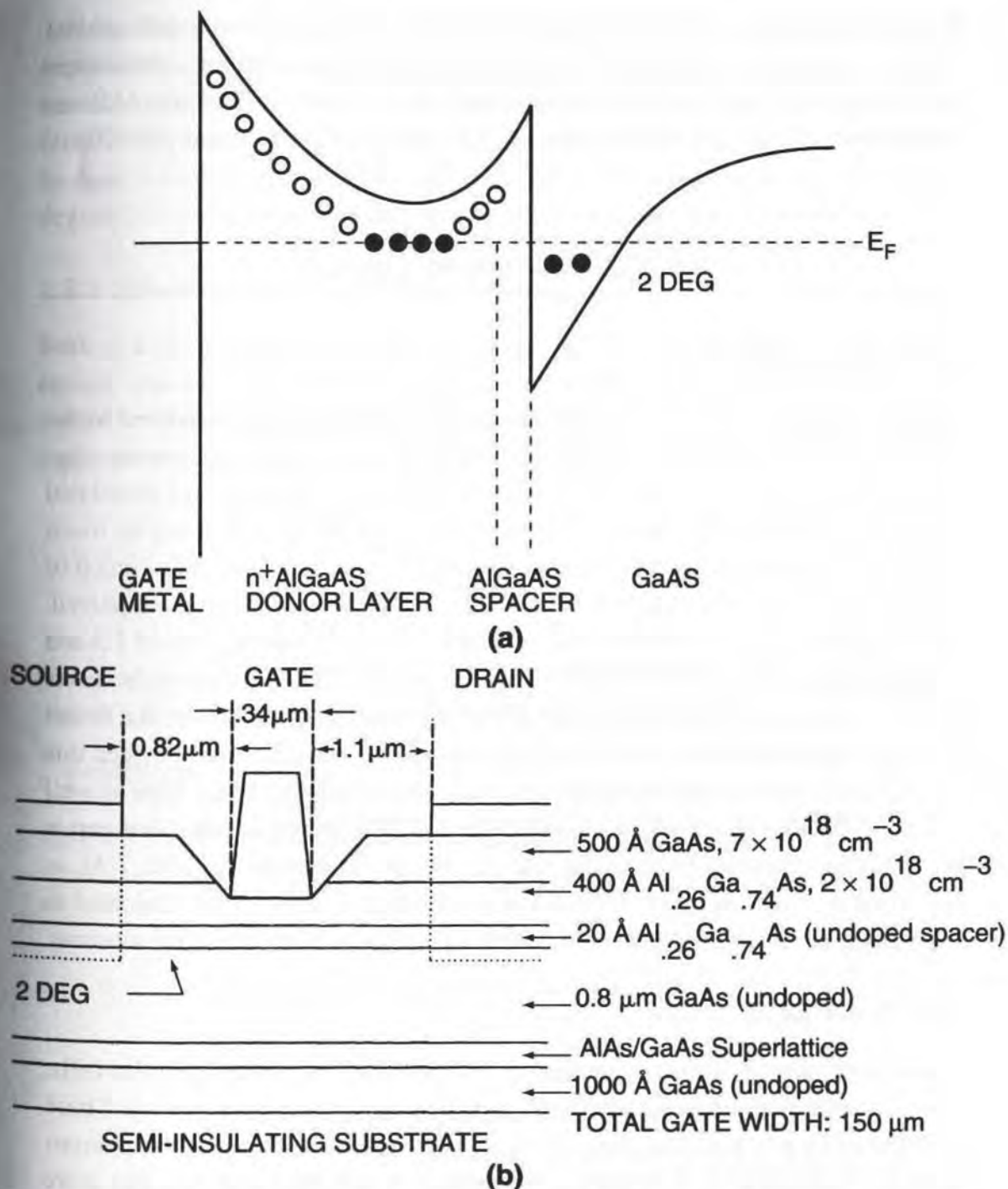


Fig. 2-24 (a) Band diagram of an n-type $\text{Al}_x\text{Ga}_{1-x}\text{As}$ -GaAs heterojunction structure. Free electrons originating in the $\text{Al}_x\text{Ga}_{1-x}\text{As}$ spill over into the quantum well on the GaAs side, creating a highly mobile 2-dimensional electron gas or carrier density. An electric field is applied normal to the paper. (b) Cross-section of a HEMT structure with dimensions to scale. The gate metal is TiPtAu, while the source and drain contacts are AuGeNi. From W.T. Anderson, K.A. Christianson, and C. Mogilestue, *Quality and Reliability Eng. Int.* **9**, 367 (1993).

MMICs composed of 0.5 μm GaAs MESFETs (with Au/Pd/Ti metallizations) tested at temperatures of 200–225°C revealed that the main failure mechanisms were diffusion related, e.g., gate sinking and ohmic contact degradation. Activation energies for failures ranged between 1.5 and 2.2 eV for the most part (20).

2.8 ELECTRO-OPTICAL DEVICES

2.8.1 Introduction

Our main concern with compound semiconductors in this book is related to the failure and reliability of electro-optical devices. The three basic phenomena capitalized upon in these devices—absorption, spontaneous emission, and stimulated emission—were already introduced in Sect. 2.2.2.4 and Fig. 2-6. Since so much is at stake in ensuring the reliability of optical communication systems, much of our concern here is with the individual active semiconductor devices employed. Today, long-distance submarine systems employ infrared wavelengths of 1.3 and 1.55 μm , where fiber transmission losses are minimal. This means that the lasers and LEDs that launch light at one end of the fiber and the photodiodes that detect it at the other end must operate at the same wavelength. To tune devices this way, semiconductors must be tailored to have the necessary energy gaps as well as a good lattice match to the substrate. Guiding the design of the band gap is Eq. 2-1. For example, by alloying Ga, As, In, and P, ternary $\text{In}_{0.47}\text{Ga}_{0.53}\text{As}$ as well as quaternary InGaAsP devices lattice-matched to InP can be fabricated to serve the needs of light sources and detectors in optical communication systems.

2.8.2 Solar Cells

A good example of devices capitalizing on absorption processes are solar cells, in which solar energy is absorbed and directly converted into electricity. These devices have a p–n junction close enough to the surface so that incident photons can easily penetrate to it. Once the junction is illuminated, carriers that were immobile in the space charge region of the prior unilluminated, or dark, junction, now move. Electrons lose energy by occupying lower conduction band levels in the n-type material; similarly, holes lose energy by floating upward to fill valence band levels in p-type material. Some of these carriers will promptly recombine. However, enough of them separate further to produce the equivalent of positive and negative battery electrodes, or a photovoltaic effect. The illuminated junction then acts like a photon-activated electrical pump, generating a current in the diode that can power an external circuit.

Reliability of terrestrial solar cells is adversely affected by environmentally generated surface deposits that reduce incident light and enhance corrosion of contact metals. Amorphous, hydrogenated silicon (a-Si:H), extensively employed in thin-film transistors for liquid crystal displays, has also been used to fabricate solar cells. In such cells, low operating temperatures and exposure to high light intensities degrade the conversion efficiency in ways that are not entirely understood.

2.8.3 PIN and Avalanche Photodiodes

Both of these diodes are normally employed as photon (absorption) detectors in optical systems and operate under reverse bias. Compound semiconductor PIN diodes are available in planar and mesa configurations (21), which are schematically depicted in Figs. 2-25a and 2-25b. These devices rely on InP and InGaAs

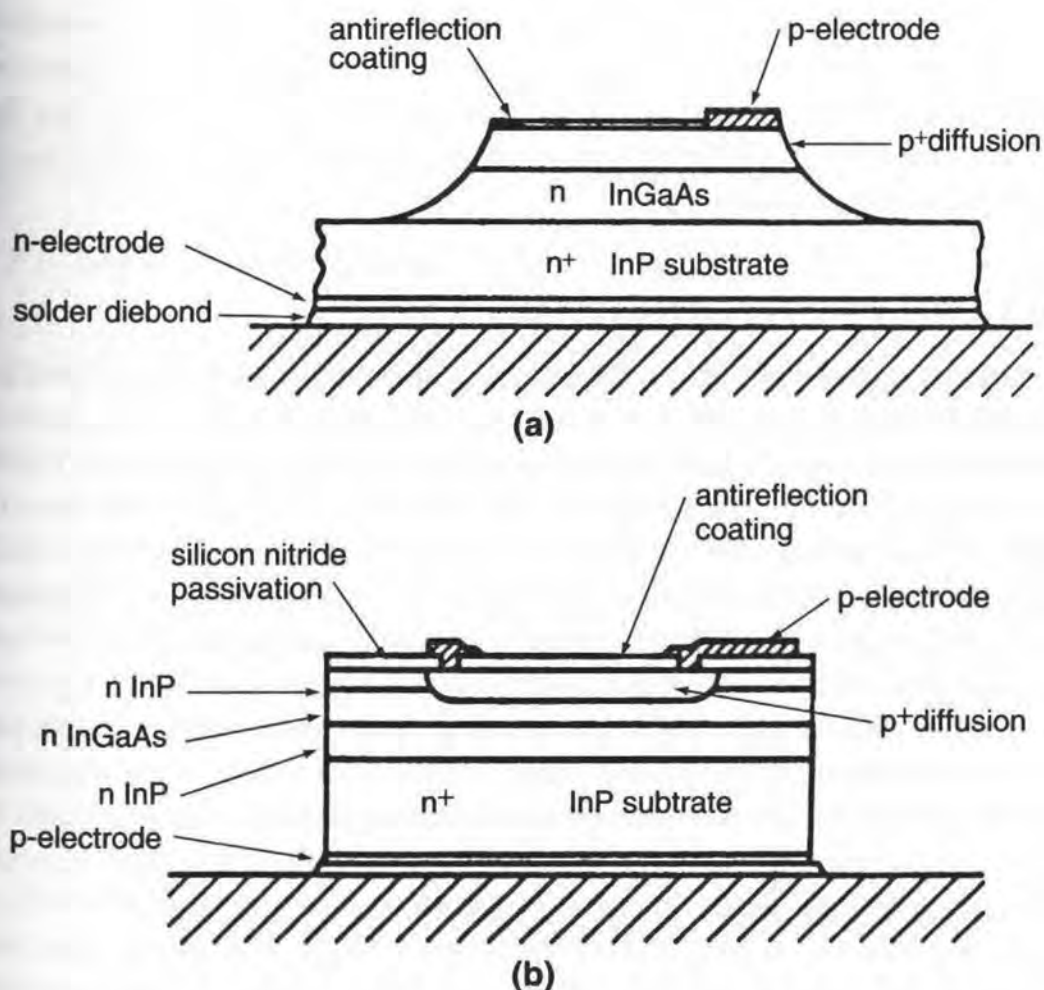


Fig. 2-25 (a) Cross-section of a III-V PIN photodiode in a mesa configuration. (b) Cross-section of a planar III-V PIN photodiode structure.

film layers to enable 1.3 to 1.5 μm wavelength operation in high reliability, long haul, and high bit-rate optical communication systems. The main objectives in designing PINs are to reduce capacitance to achieve high bandwidth, eliminate noise, and minimize reverse-bias dark current, which limits receiver sensitivity. Easy-to-fabricate mesa structures enable the first objective to be met, but such devices exhibit dark current instabilities. Planar PIN structures have a passivation layer that protects the junction, and a p-doped InP cap that minimizes leakage current. As a result, planar PINs are inherently more reliable than mesa devices.

Avalanche photodiodes (APD) are similar to planar III–V semiconductor PINs but operate at much higher voltages to obtain multiplication of the photo-generated carriers. Light is absorbed in the InGaAs layer, and holes are swept to an InP junction where the avalanche multiplication of carriers initiates. The separation of absorption and multiplication regions reduces the junction leakage current. Reliability problems arise from high electric fields, which induce surface charge motion (22). Other failure mechanisms that lead to increased dark currents stem from contact metal penetration of the junction region and localized breakdown of the junction edge. Although lifetimes of APDs are shorter than for PIN devices, the high failure activation energies of 1 eV result in reliable operation under normal conditions.

2.8.4 Light Emitting Diodes (LEDs)

When these diodes are forward biased, they spontaneously emit light as conduction-band electrons recombine with valence-band holes. This phenomenon of injection electroluminescence has been exploited in LEDs to produce the familiar colored numbers and letters in displays. LEDs are also a viable alternative to lasers in some short to medium distance optical communications systems, where their lower light output is not a disadvantage. Such an application is illustrated in Fig. 2-26, where the LED, composed of InGaAsP lattice-matched to the InP substrate, launches 1.3 μm infrared light down the optical fiber.

Degradation in these devices stems from both the semiconductor materials and the contact metals (23, 24). Typical p-side metallizations employed are Au–Zn or Au/Pt/Ti. Devices using the latter trimetal contacts generally exhibit less loss in light output during aging at temperatures of 120–200°C. For the Au–Zn LEDs the activation energy for failure has been determined to be in the range 0.75 to 0.85 eV. At high light power, degradation is primarily caused by dark spot and dark line defect formation, a mode that also afflicts lasers. Precipitation of host atoms and the familiar migration (sinking) of contact metal into the semiconductor have been identified as degradation mechanisms. Nevertheless, edge-emitting diodes

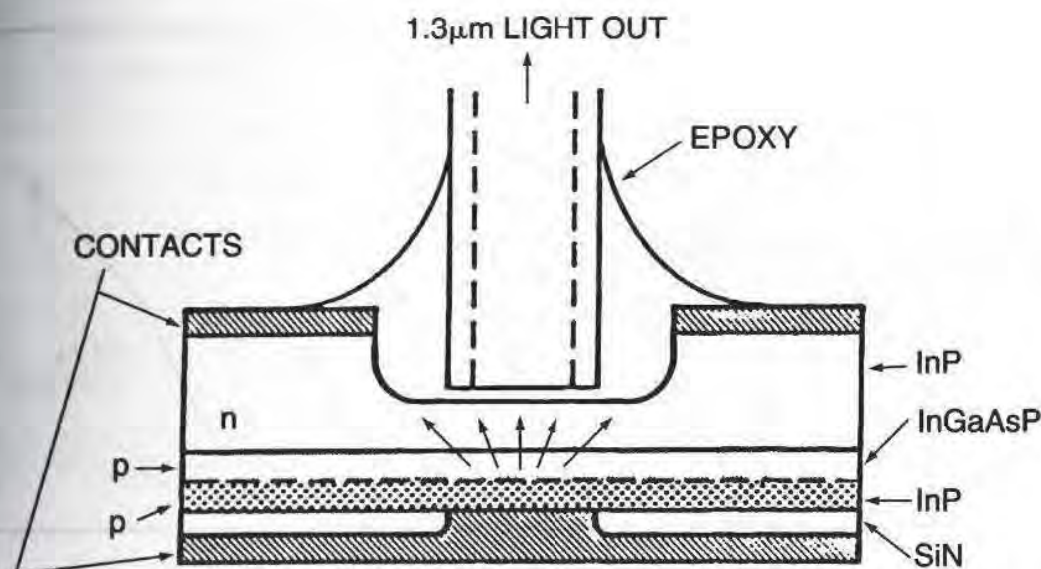


Fig. 2-26 Schematic of a surface-emitting InGaAsP light-emitting diode lattice-matched to an InP substrate. The $1.3\ \mu\text{m}$ infrared light emitted enters an optical fiber.

(ELED) are considered to be highly reliable devices, especially at low power levels, where lifetimes of $\sim 10^7$ hours at 25°C have been projected.

2.8.5 Semiconductor Lasers

Due to the many critical laser applications there has been far more published information on the reliability of lasers than of other opto-electronic devices. While light is emitted from LED junctions in all directions, intense, directed beams exit lasers. In order for spontaneous emission of photons and laser action to occur, however, it is first necessary to have a population inversion of carriers. This is accomplished in a semiconductor diode structure where both p and n regions are so heavily doped that the Fermi levels lie inside the valence and conduction bands. Under conditions of high forward bias, large electron and hole concentrations exist within respective bands that now overlap in the vicinity of the junction. A carrier population inversion is created, and stimulated emission occurs. Parallel facets are cleaved and polished, forming the Fabry-Perot cavity in which the photons multiply, while the other faces are roughened to discourage lasing in these directions.

Emitted laser light has a wavelength corresponding to the band-gap energy (Eq. 2-1) and exhibits the so-called $L-i$ (L = optical power, i = current) output characteristics schematically indicated in Fig. 2-27. Up to a critical threshold current density (j_{th}), the light output is very low, but beyond j_{th} it rises rapidly. Higher operating temperatures raise j_{th} and lead to lower light outputs. Laser

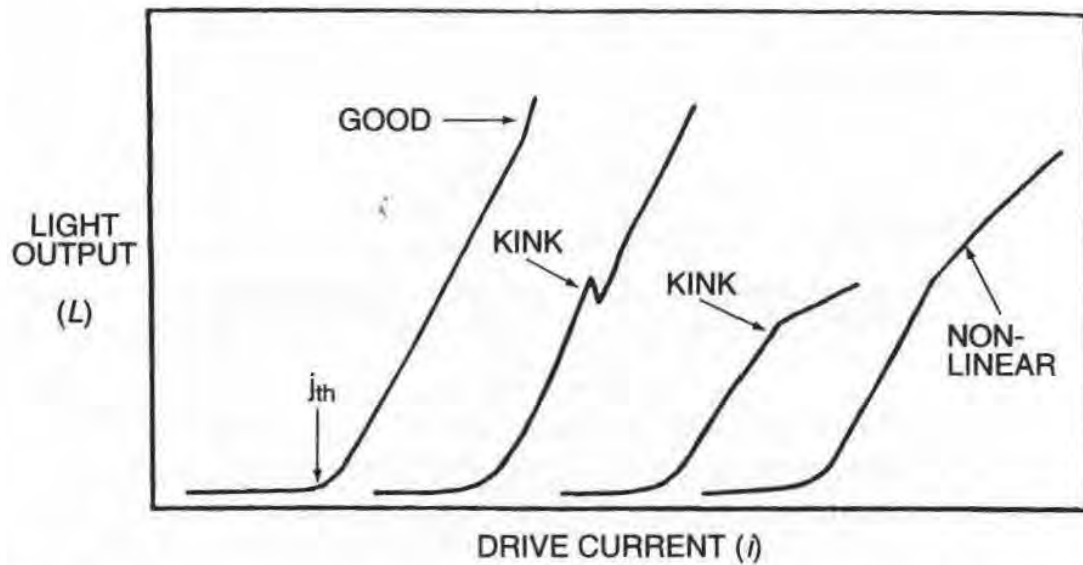
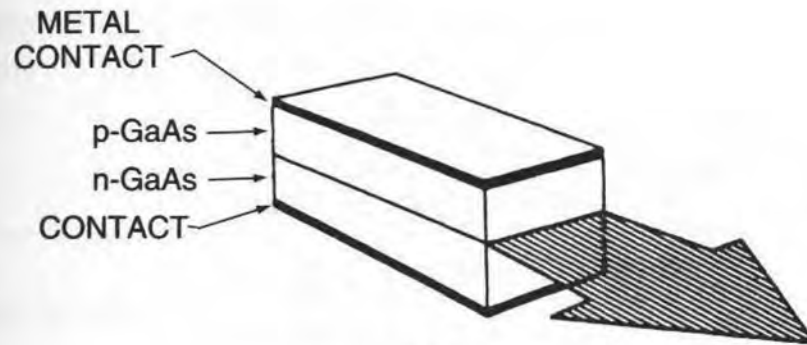


Fig. 2-27 Assorted laser light output current characteristics depicted schematically. From reference 25.

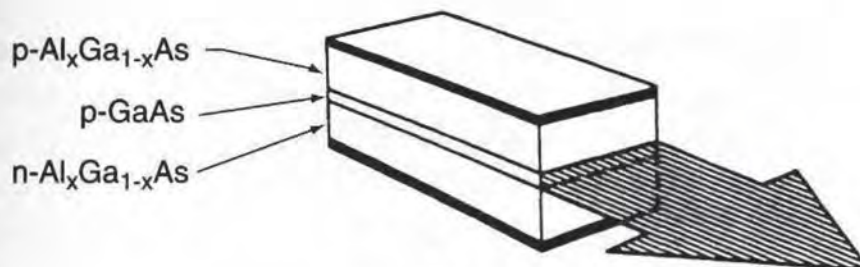
diodes displaying the kinked and nonlinear characteristics shown in Fig. 2-27 are considered to be faulty (25). Peaked kinks are often, but not always, indicative of internal defects and damage, while internal reflections frequently, but not always, cause the second type of kink characteristics. Kinks are also associated with higher-order modes of laser emission and changes in j_{th} .

The simplest of all laser structures, the *homojunction* configuration, contains the same semiconductor material, e.g., GaAs, on both sides of the junction. Good only as a pedagogical tool, such lasers require enormous threshold current densities ($\sim 50 \text{ kA/cm}^2$ at 300K and $\sim 5 \text{ kA/cm}^2$ at 77K) to initiate lasing. The severe reliability problems implied by Joule heating thus rule out continuous homojunction laser operation at room temperature. Furthermore, carriers can wander from the active region and recombine, resulting in a broadened emission. Output light stems from a considerably narrower active region in *double-heterojunction* lasers. First, the carriers are confined by the two heterojunction GaAs and $\text{Al}_x\text{Ga}_{1-x}\text{As}$ barriers. Second, light is confined to the active region by virtue of the higher index of refraction of GaAs relative to AlGaAs. Both effects lower the lasing threshold an order of magnitude, while narrow stripe contacts ($20 \mu\text{m}$ wide) reduce j_{th} further, to the point where reliable continuous operation can occur at room temperature. The evolution of these GaAs-based lasers, designed to operate in the range $0.8\text{--}0.9 \mu\text{m}$, is depicted in Fig. 2-28.

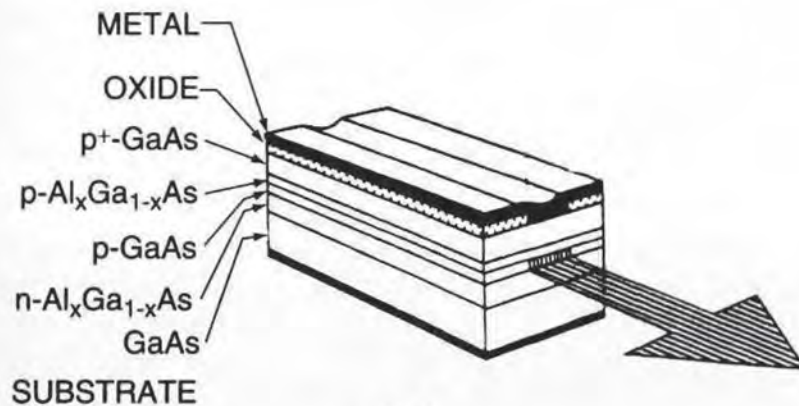
Popular laser geometries for optical communication systems are shown in Fig. 2-29 and employ InGaAsP semiconductor materials (25). Known as BH or buried-heterojunction lasers, the light generated in the buried active regions is



(a)



(b)



(c)

Fig. 2-28 Evolution of GaAs-based lasers, designed to operate in the range 0.8–0.9 μm . From S.M. Sze, *Semiconductor Devices, Physics and Technology*, J. Wiley and Sons, New York (1985), Copyright by Bell Telephone Laboratories, Inc.

strongly guided in these devices. Other features include low threshold currents and high quantum efficiencies. Instead of using only the reflective facets for controlling light output, an extended diffraction grating is incorporated in so-called distributed feedback (DFB) lasers for this purpose. In this way a single laser wavelength from the multi peaked laser output can be selected and amplified.

A number of failure modes commonly observed in lasers include damage to facets, dark line and dark spot defect formation, aging effects, electrostatic

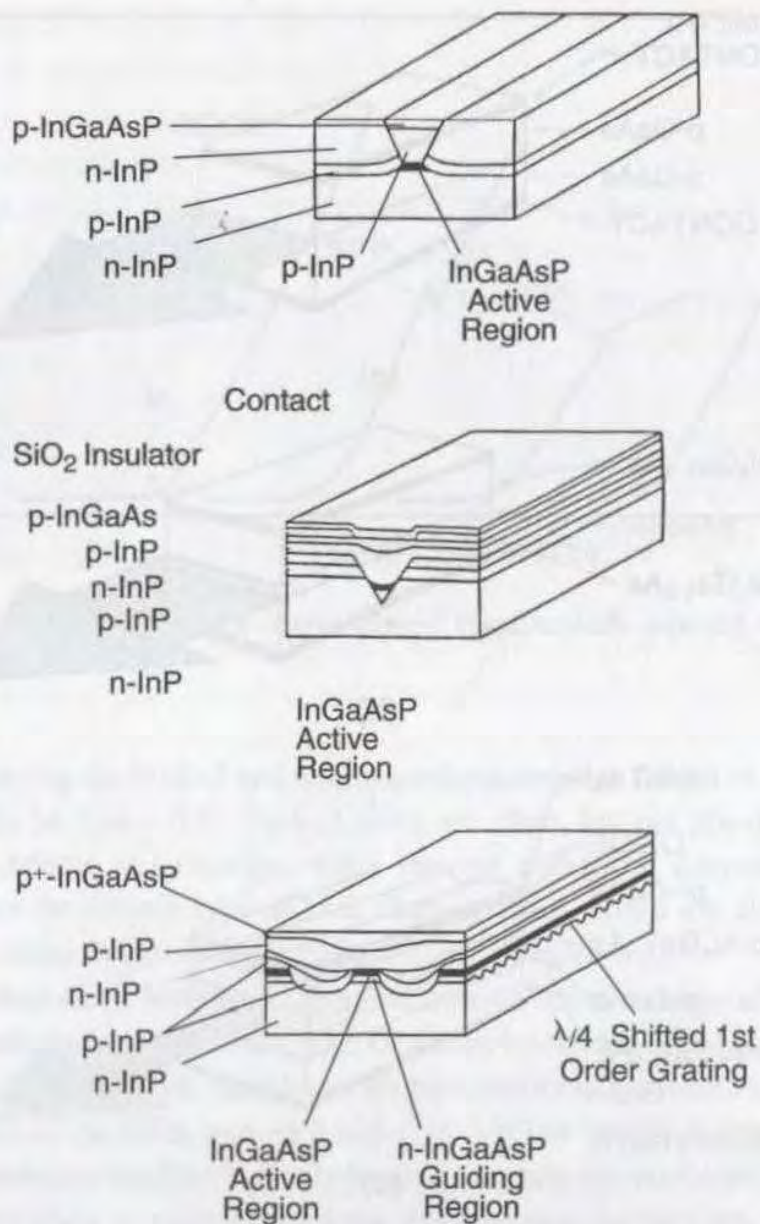


Fig. 2-29 Popular laser geometries for optical communication systems. *Top:* Etched mesa buried heterostructure (EMBH). *Middle:* Channeled substrate buried heterostructure (CSBH). *Bottom:* Double channel planar buried heterostructure distributed-feedback. From reference 25.

discharge, and current-confinement problems. The latter are manifested as a drop in laser efficiency because current leaks or is shunted away from the active region. Poor design, processing, and power surges are often the cause of current-confinement problems, to which some laser structures are more prone than others. Interestingly, InGaAsP lasers have a higher immunity to laser degradation effects than AlGaAs/GaAs-type lasers.

Detailed mechanisms for the failure of lasers and other electro-optical devices and components will be presented in Chapter 10.

2.9 PROCESSING—THE CHIP LEVEL

2.9.1 Introduction

It is much easier to draw schematics of devices and IC circuit diagrams on paper or with the aid of computers than to realize them in physical terms. The production of devices requires different electrical materials to be shaped or patterned and creatively merged to form the active devices and passive components that must then be interconnected, insulated, and finally packaged. These issues and demands are at the core of a very broad interdisciplinary manufacturing technology that designs, fabricates, and packages solid-state devices and circuits.

Much of the glamour associated with semiconductor manufacturing occurs on the Si chip level, and that is what we shall address here. To appreciate what is involved, consider the state-of-the-art CMOS technology employed in the fabrication of advanced microprocessor chips. At the base of the stunning image of Fig. 2-30, beneath the skeleton of the exposed interconnects, contacts, and first aluminum wire level, are the CMOS transistors. These MOSFETs have effective channel lengths of $0.25\text{ }\mu\text{m}$ and gate-oxide thicknesses of 7.7 nm (26).

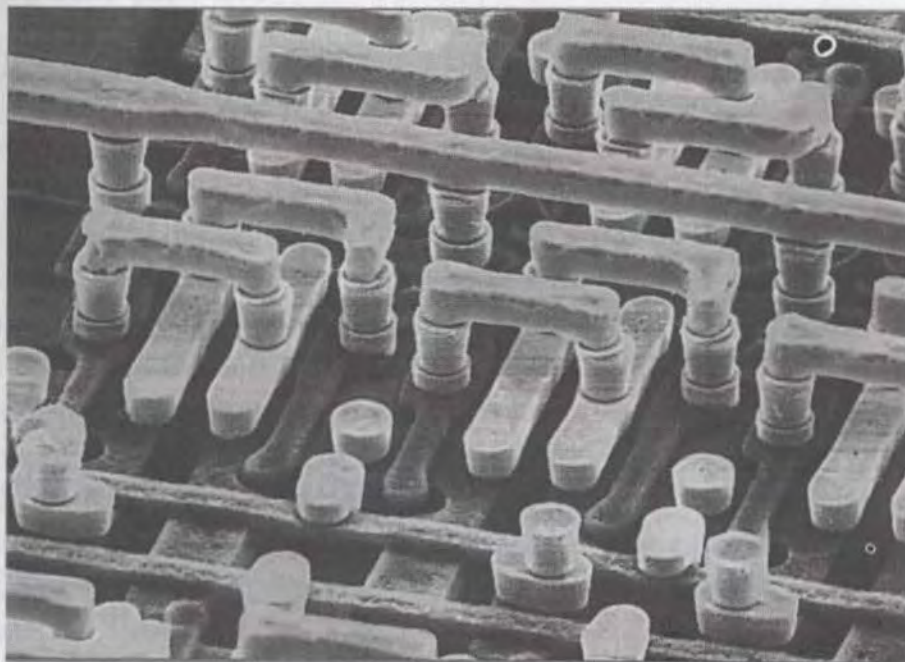


Fig. 2-30 Scanning electron microscope image of the local interconnect and first-level aluminum wires in an SRAM cell. From reference 26.

Issues related to the processing and fabrication of both Si and compound semiconductor devices have been very widely treated in the literature and form the subject of many excellent books (6–10). In this section some of the more important individual process steps in the fabrication of silicon IC chips are briefly introduced. The intent is to convey a sense of the physical and chemical principles involved, and how potential defects may be introduced into the product. Processes discussed are number-keyed to the schematic of the MOSFET shown in Fig. 2-31. In the following chapter specific processing defects will be addressed in greater detail.

2.9.2 Silicon Crystal Growth (1)

When defects or reliability problems surface in any manufactured product, our first reaction is to suspect the quality of the starting materials. Wafers are the starting materials, but they are sliced from large single crystals. Therefore, growing crystals is the first, and in many ways a critical step, in ensuring product quality. Bulk single crystals of silicon are grown by melting a Si charge in a SiO_2 crucible, inserting a small cooled single-crystal seed into it, and pulling out the large solidified crystal. This is the Czochralski technique for growing “CZ” crystals. Successive solidification occurs amid complex heat and mass transfer processes between and among the growing single crystal, molten Si, and surrounding inert gas ambient. Even more complex is the simultaneous convective fluid motion that disperses melt dopants and impurities. This fluid flow is dependent on melt temperature and relative speed of rotation between crystal and melt as the former is being pulled.

Simultaneously, compositional change occurs in both the melt and the crystal. To see why, look at almost any binary Si-dopant (P, B, As) or impurity phase dia-

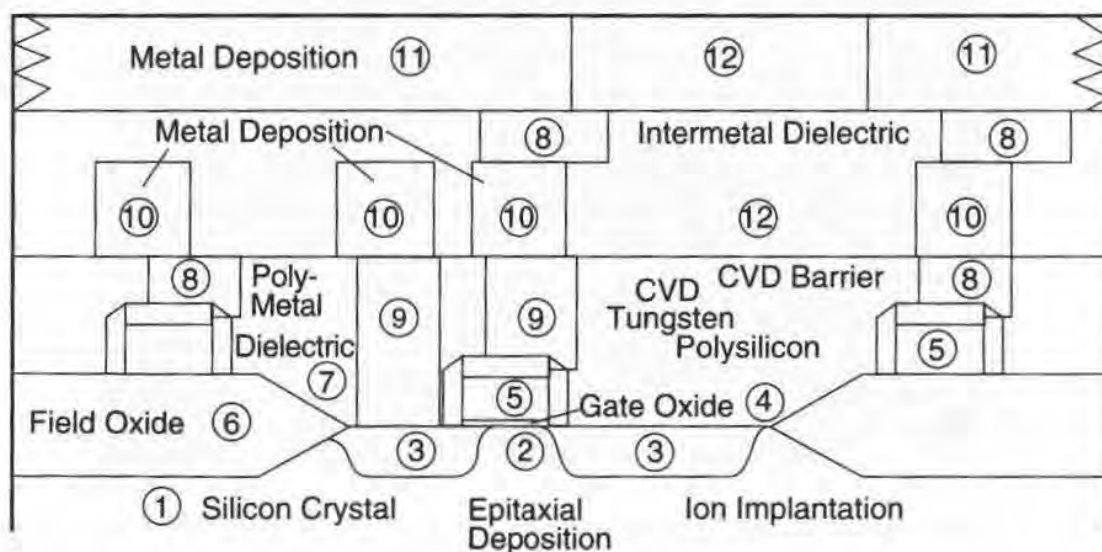


Fig. 2-31 Schematic of the MOSFET indicating key processing steps.

gram. Note that the liquidus and solidus phase boundary lines drop for small solute additions. Equilibrium at the crystal–melt growth interface implies that the solid is always purer than the liquid. Since the effect is cumulative during directional solidification, axial gradients in composition are predicted (27). Grown crystals not only exhibit varying concentration profiles of intentionally added dopants, but contain unwanted elements like oxygen. But it is the transition metal impurities that are viewed with particular alarm; they cause the most severe problems, and great efforts are made to eliminate them or neutralize their influence. In addition, defect swirl patterns are not infrequently observed on planes sliced perpendicular to the crystal axis. Furthermore, surface striations that are several microns wide and alternately dopant rich and dopant poor sometimes decorate the length of bulk Si and compound semiconductor single crystals. Both features apparently stem from temperature fluctuations at the growth interface and complex convective flow patterns that cause a periodic partitioning of solute. From the standpoint of the MOSFET, variations in dopant level across the plane of a wafer must be minimized to ensure uniform device processing and chip performance.

Following growth, the crystals are shaped round by grinding and then sawed and polished to produce wafers. Presently, integrated circuit production lines process 15 to 20 cm diameter Si wafers, but huge 30 cm diameter wafers are expected to become the world standard shortly after the turn of the twenty-first century.

2.9.3 Epitaxy (2)

This important term refers to the ordered crystallographic registry of a single crystal film deposit on an underlying crystalline substrate template. If film and substrate are the same (Si on Si) we speak of *homoeptitaxy*. In electro-optical devices they are different (InGaAsP on InP), and *heteroeptitaxy* is the appropriate term. The yield of early thin-film bipolar transistors fabricated within homoeptitaxially deposited Si was dramatically improved relative to those produced within the original, or uncoated, Si wafers. Chemical vapor deposition (CVD) or some variant of it is used to deposit epitaxial semiconductor films on wafer substrates. The process typically entails reaction of a chlorinated silane, e.g., SiCl_4 , and H_2 at temperatures of $\sim 1100^\circ\text{C}$ and a pressure of ~ 1 atm. Epitaxial processing in CMOS technology helps assure device isolation and minimizes junction leakage currents.

Critical to insuring single crystal formation is a combination of a low film nucleation rate and a high film growth rate. If many nuclei form on the substrate, a worthless polycrystalline film deposits. A low gas supersaturation and deposition rate will create few nuclei. On the other hand, high substrate temperatures enhance diffusion rates and facilitate atomic incorporation into lattice sites. This promotes the early lateral extension of the single crystal film, and growth perfection as it thickens. Epitaxial Si films are purer and frequently more defect-free

than the substrate wafer; furthermore, they can be doped independently of it. Nevertheless, they do contain the assorted point, line, and surface defects noted later in Fig. 3-15 that occasionally cause device malfunction.

2.9.4 Ion Implantation (3)

For high-performance devices with shallow junction depths, ion implantation, rather than diffusion, is the preferred method of doping. Unlike diffusion, where the surface dopant concentration is higher than in the crystal interior, the reverse is true in ion implantation. The subsurface wafer penetration of dopant ions is due to energies of ~50–200 keV acquired when they are accelerated in the high electric field of the implanter. Possessing energies thousands of times larger than the binding energy of lattice atoms, the dopant ions wreak havoc as they are embedded into Si. At first a lightning-like ion current excites the atomic electrons and ionizes atoms, absorbing energy in the process. When the dopants lose enough energy, they slow sufficiently and then begin to violently collide with Si nuclei, knocking many off their lattice sites. Each dopant ion executes a different odyssey and comes to rest at a different distance from the wafer surface. Spatially, dopant concentrations are statistically distributed as a Gaussian function both normal to the surface and transverse to it. The depth of the peak and spread of the Gaussian depends on the dopant and the nature of the matrix it passes through, e.g., Si, SiO₂, SiN, photoresist. Ion beam damage to the silicon matrix creates disordered and even amorphous regions. These must be repaired by rapid thermal annealing at ~900°C, where atoms can diffuse sufficiently to restore lattice crystallinity and full electrical activity to the dopants.

In modern IC chips some 15 to 20 or more ion implantation steps are carried out to dope large volumes (tubs), small regions (sources and drains), and to adjust the threshold voltage of field effect transistors. This latter operation, which controllably alters the interface charge (Q_i in Eq. 2-6) at the Si–SiO₂ boundary, is quite critical and requires precise ion doses and spatial distributions. Tailoring dopant profiles by ion implantation has additionally enabled minimization of hot-electron damage and latchup effects, as well as gettering of heavy metal (Fe) contaminants. In the latter two cases, high-energy implants are required to drive dopants deep beneath the surface layer of devices.

Ion implanters are one of the most complex pieces of microelectronics processing equipment; they are also potential and actual sources of contamination and defects. Ideally, a chemically pure, monoenergetic beam is desired. In a practical sense one may imagine defects arising from variations in ion beam energy, depth, lateral spatial distribution and ion dose, as well as from the presence of chemical contaminants. Examples are given in Chapter 3.

All of the processes to this point have been concerned with the underlying Si. Above the Si surface are assorted grown and deposited dielectric and conducting films that we now consider.

2.9.5 Gate Oxide (4)

Today's ULSI technology is due in large part to the excellent properties of thermally grown gate oxides. This critical SiO_2 film, the smallest fabricated dimension in MOS devices, has been the focus of much study and reliability testing. Currently, less than 10 nm thick, it is made by controllably oxidizing silicon wafers in dry O_2 . The simple reaction, $\text{Si} + \text{O}_2 \rightarrow \text{SiO}_2$, at temperatures between 900 and 1100°C, produces a dense amorphous film. Thick oxide growth has traditionally been modeled by linear kinetics at short time (i.e., $d_o \approx t$) and by parabolic kinetics at long time, i.e., $d_o \approx t^{1/2}$, where d_o is the oxide thickness and t is the time. These formulas may not hold for the very thin oxides now grown.

Growth in the presence of N_2O is a recent advance that has improved the quality of gate oxides. Nitrided oxides are apparently more resistant to (charge) trap generation and to plasma-induced charging, give rise to smaller electric-field-induced leakage currents, and are generally longer lived than conventional oxides (28, 29). Oxide charging during operation, and the associated reliability problems of breakdown and hot electron degradation are topics discussed in Chapter 6.

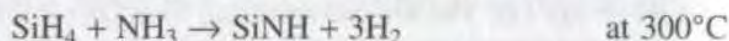
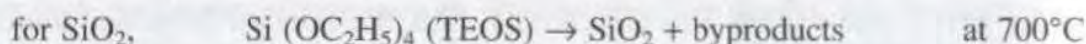
2.9.6 Polysilicon (5)

The gate oxide is contacted by a polysilicon electrode. Long ago, polysilicon replaced the much more unreliable aluminum electrodes that caused early dielectric breakdown in underlying oxides. Polysilicon films are deposited in low-pressure CVD reactors through the reduction of silane by hydrogen. Below ~590°C amorphous films form, while above ~610°C polycrystalline films deposit. This narrow temperature range defines a useful processing window. At higher temperatures, gas phase reactions result in rough, loosely adhering deposits and silane depletion that causes poor film uniformity. Impracticably low deposition rates occur at lower temperatures.

2.9.7 Deposited and Etched Dielectric Films (6,7,12)

Since all devices and interconnections support voltages and carry current while sitting on a conducting Si wafer, strategically placed dielectrics are needed to isolate them electrically. Aside from the critical (grown) gate oxide, an assortment of other insulator films are *deposited*. They include SiO_2 , silicon nitride, and phospho- and boro-silicate glasses produced by CVD methods, often with the assistance of a plasma discharge. The dielectric films must then be etched to

expose windows to semiconductor contacts or to confine subsequently deposited metal films and vias (see next section). After microelectronic chips are fabricated they are usually encapsulated by a silicon nitride film, an effective barrier that prevents moisture and alkali ion (Na^+) penetration. All of the above dielectrics are amorphous in structure and of the order of $1\text{ }\mu\text{m}$ thick. Typical reactions with precursor gases (on the left) and product films (on the right) include:



for phosphosilicate glass,



In general, dielectric film deposition steps are carried out at successively lower temperatures so as not to alter the prior processed and doped materials.

The composition, structure, and properties of such CVD films vary widely with deposition conditions. Silicon nitride is an example. At very high temperatures dense, crystalline Si_3N_4 forms, but at lower temperatures considerable hydrogen is incorporated, particularly during plasma assisted deposition. Differences in dielectric strength, internal stress, and resistance to permeation of water and ionic impurities occur, but generally not to the point of compromising reliability. High phosphorous or arsenic glass compositions also corrode aluminum lines when moisture is present. More significant are the problems associated with surface coverage by dielectric films and maintaining a planar terrain for subsequent metal depositions. Incomplete filling of deep trenches and regions between metal lines can lead to structural voids and occasional short circuiting.

Maintaining a planar topography is extremely important if subsequent lithography patterning errors and defects in deposited structures are to be avoided. The relatively new technique of chemical-mechanical polishing (CMP) has enabled the production of planar interfaces between successive metal and intermetal dielectric levels.

2.9.8 Metallization (8–11)

There is a need to contact and interconnect all of the doped semiconductor electrodes, i.e., source, drain, and gate, to other components and devices on the chip. Assorted metal and alloy conductors are employed for contact structures including Al and Al–Si, PtSi, Ti–W, Pd_2Si , TiSi_2 . Interconnections, on the other hand, are invariably composed of Al thin-film alloys with dilute Cu or Cu–Si additions. In high-density ICs, efficient packing of devices on the chip real estate requires

placing interconnect stripes at successive metal levels (30) commonly denoted by M1, M2, M3, etc., as shown in Fig. 2-32. The highest level is composed of bus-bars that carry current to the chip contact pads. Vias in the form of studs or plugs are like staircases that connect neighboring floors, or levels, of interconnects. Tungsten plugs or vias are often deposited by CVD methods to vertically link

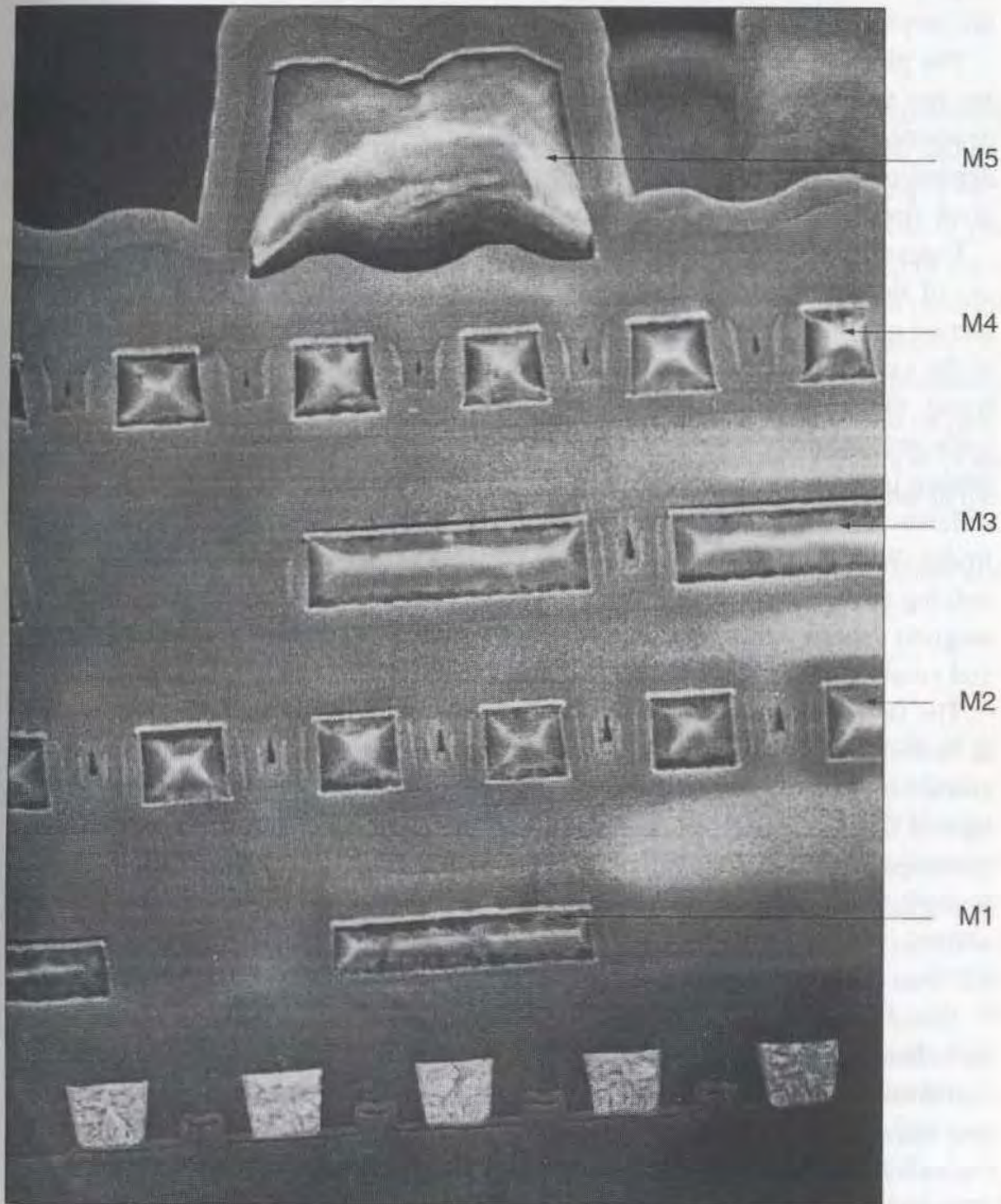


Fig. 2-32 Scanning electron microscope image of five levels of metallization with a local interconnect. From reference 30.

adjacent interconnect levels. As shown in Fig. 2-31, these vias are surrounded by TiN barrier films that prevent mobile ions from reaching sensitive junctions and reactive gases from etching exposed oxides. The above metals possess the best collective basket of desirable properties, including 1. high electrical conductivity; 2. low contact resistance to Si; 3. little tendency to chemically react with or interdiffuse in Si; 4. resistance to corrosion and environmental degradation; 5. electromigration resistance; 6. ease of deposition; 7. compatibility with other materials and steps in the fabrication process.

The physical vapor deposition (PVD) techniques of *evaporation* and *sputtering* are the means employed to deposit metal films. Sequential atom-by-atom deposition results in film nucleation and subsequent growth. Unlike epitaxial deposition, both methods yield polycrystalline films typically having submicron grain sizes, due in part to substrates that are often amorphous (e.g., SiO₂).

Evaporation is the simpler of these metallization methods and involves heating of the source metal until it evaporates at appreciable rates. The process is carried out in high vacuum while wafer substrates, carefully positioned relative to the evaporation source, are rotated to yield uniform metal coverage. In sputtering, the metal or alloy to be deposited is fashioned into a cathode plate electrode and generally placed parallel to the wafers (anode) within a chamber containing inert sputtering gas. At low pressures less than ~5 mTorr, and potential differences of a few kilovolts dc, a glow discharge is sustained between electrodes. Within the discharge, energetic positive gas ions bombard the target, dislodging atoms that then deposit on the wafers as a film. In magnetron sputtering, magnets placed behind the cathode confine electrons to hop around the target and create a richer plasma that increases film deposition rates.

The two metallization processes compete in many applications, and not only in semiconductor technology, but other industries as well. Issues that must be considered in selecting the best one for a particular application include 1. adhesion to the substrate; 2. film stress; 3. stoichiometry of alloys films; 4. uniform coverage of the topography; and 5. electrical damage to substrate oxides. Sputtering is preferred in most semiconductor applications because stoichiometry of alloys is maintained and step coverage is usually superior. However, evaporated Al films appear to be more resistant to electromigration damage.

Many reliability problems are associated with contact and interconnection metallizations. They include changes in contact resistance due to interdiffusion reactions between Al and Si, electromigration-induced degradation of contacts and interconnections, and general corrosion damage. Contacts between Si and Al, two indispensable elements in microelectronics, are particularly troublesome. This results from mutual atomic interdiffusion (Section 5.4.2) and is remedied through alloying (e.g., Al-1%Si) or the use of diffusion barriers (e.g., metal sili-

cides, TiW). These must be imposed between the silicon and Al interconnections to prevent reaction. Other problems stem from stress. When surrounded on all sides by oxide, as interconnects are in the various metallization levels, tensile stresses develop in the metal. The reason, as we shall see in the next chapter, is due to the larger thermal expansion coefficient of Al relative to SiO_2 . These thermal stresses are sometimes sufficiently large to generate voids that open interconnects.

2.9.9 Plasma Etching

Plasma processes are widely employed in the fabrication of semiconductor devices. Examples include plasma-enhanced CVD, photoresist removal, and the etching of polysilicon, dielectric, and metal thin films. Within the plasma a very reactive environment consisting of high energy electrons, ions, and precursor gas free radicals is created by applied voltages of a few kV. This combination of energetic reactants speeds up otherwise sluggish vapor deposition or material removal (etching) reactions, and effectively makes them proceed at lower temperatures. In our MOSFET, the need to etch contacts and vias in SiO_2 is met by radio-frequency (RF) plasmas employing process gases containing fluorine, carbon, and oxygen. Minimum dimensions, high aspect ratio features, and selectivity of the etch are necessary attributes of the etching process. The last attribute requires that neighboring polysilicon, metal, or silicides be immune to etching attack.

Damage to materials during plasma etching includes the buildup of charge and dielectric wearout due to bond breaking and charge trapping.

2.9.10 Lithography

Perhaps the most critical process in microelectronics is lithography, a technology that embodies a collection of critical steps. Lithography involves the use of exposure tools, masks, resist materials, and etching processes to enable such blanket-area processing steps as film deposition, etching, diffusion, and ion implantation to effectively occur only in preselected areas of a wafer surface. While there are obvious parallels to photography, instead of producing flat images on paper, a three-dimensional relief topography is created in materials like Si, SiO_2 , silicon nitride, aluminum, and photoresist. After a mask (reticle) is generated with the desired geometric features, lithographic patterning involves spinning a photoresist film on the wafer surface. As the name implies, photoresists are sensitive to photons and resistant to chemical attack by liquid or gaseous etchants. Some are also sensitive to electrons and X-rays. In general, they consist of specially compounded polymers and display two broad types of behavior. When exposed to light, the cross-linked polymer chains of *positive resists* undergo scission, or cut-

ting, rendering them more soluble in a developer, while unexposed areas resist chemical attack. Under the same stimulus, *negative resist* molecules polymerize, while unexposed regions are soluble and dissolve in the developer.

If, for example, it is desired to dope silicon, a blanket film of SiO_2 is first deposited and then covered with a photoresist layer. After exposure and development, the unprotected SiO_2 film is removed by etching, and the photoresist is stripped away. This leaves the bare Si surface geometrically patterned by the remaining SiO_2 film mask. Now the exposed Si regions may be ion implanted. Similar steps must be repeated at many different depths, or levels, on the wafer as actual integrated circuits unfold. Alignment of successive levels, e.g., contact metal over a contact hole, is absolutely critical so that correct registry is maintained simultaneously over millions of device features on large wafers. The magnitude of the problem can be appreciated when it is understood that 10 to 20 lithography levels are required in advanced IC chips.

Poor product yields have often been associated with lithography defects. These arise predominantly from mask pattern defects that are discussed in Sect. 3.4.2.4. Metallization patterns defined after etching may then, for example, contain either missing or added metal, which results in open or shorted conductors, respectively. Mask defects can be more readily tolerated by the expedient of employing $5\times$ or $10\times$ reduction stepper tools that project the patterns onto photoresists. The $1\times$ systems do not offer this opportunity and therefore require more defect-free masks.

EXERCISES

- 2-1 What are some differences between materials used for electrically and mechanically functional applications?
- 2-2 Metals are the best conductors of electricity and heat, reflect light most efficiently, and are the most malleable of materials. Explain these attributes of metals in terms of their electronic and atomic structure.
- 2-3 Copper wiring in a telephone switching station was exposed to a fire. After the fire the electrical resistivity of the wiring was measured. Some wiring was observed to have a lower resistivity, while other wiring exhibited an increase in resistivity.
 - a. Describe a scenario that would result in a resistivity decrease.
 - b. Describe a scenario that would result in a resistivity increase.

- 2-4 A good dielectric is a good insulator, but a good insulator is not necessarily a good dielectric. Explain why.
- 2-5 a. Rectifiers conduct large currents in one direction. Metals in contact also conduct large currents, but in both directions. Why can't rectifiers be fabricated using metals?
b. An electric field is built in at a semiconductor p-n junction. Insulators can support applied electric fields. Why can't useful rectifiers be fabricated using insulators?
c. What properties make semiconductors indispensable for junction devices.
- 2-6 Electronic materials that will probably play important roles in the future are GaN for light sources, $\text{YBa}_2\text{Cu}_3\text{O}_7$ for interconnections, and titanates for capacitors. Approximately where on the scale of conductivities in Fig. 2-1 do these materials belong?
- 2-7 The p-type regions of two discrete diodes are connected in common. Why does the resultant n-p-n structure not behave like a transistor?
- 2-8 Distinguish between a. bonds and bands, b. thermal and impurity scattering, c. spontaneous and stimulated emission, d. Fermi energy and energy gap, e. Direct and indirect band-gap semiconductors, f. electronic and opto-electronic devices, g. metal-semiconductor and p-n junction diodes, h. ohmic and Schottky contacts, i. MESFETs and HEMTs, j. bipolar and MOS transistors.
- 2-9 High local electric fields promote avalanche breakdown effects in both semiconductors and dielectrics. Such breakdown is usually harmless in semiconductors but can be destructive in dielectrics. Why? Electron-avalanche effects do not occur in metals. Why?
- 2-10 High-power semiconductor diodes are usually fabricated containing a heat sink. To understand why, sketch the effect of high operating temperatures on diode current-voltage characteristics.
- 2-11 When ohmic and Schottky contacts suffer electrical degradation, some alteration of the metal-semiconductor band diagram is expected. Indicate possible band-diagram changes in the vicinity of a contact and explain why they occur.
- 2-12 Boeing Aerospace conducted reliability studies on some 25 million semiconductor devices and found that bipolar power devices failed significantly less often than NMOS and CMOS integrated circuits. What are possible reasons for this observation?

- 2-13 In a real, or nonideal, MOS capacitor structure, the work functions of the metal and semiconductor differ. Sketch the band structures of an unbiased metal-SiO₂-p-type Si capacitor if $\phi_M > \phi_{Si}$.
- Illustrate depletion, accumulation and inversion for this structure.
 - Sketch the band structures for depletion, accumulation, and inversion in an ideal metal-SiO₂-n-type Si capacitor structure.
- 2-14 One strategy to minimize the probability of latchup in CMOS devices is to fabricate them within a silicon film grown epitaxially on an insulator (SOI). Illustrate how SOI structures suppress latchup.
- 2-15 Distinguish between the functions of a. Microprocessor and memory chips, b. DRAMs and SRAMs, c. Charge-coupled devices and DRAMs, 4. Gate capacitor and trench capacitor.
- 2-16 When zero bias is applied in the HEMT structure, the conduction band edge in the GaAs lies below the Fermi level, implying a large density of the 2-D electron gas.
- What do you expect to happen when the gate is biased negatively?
 - Draw band diagrams for both zero and negative bias.
- 2-17 When compared to silicon device technology, compound-semiconductor laser production involves may fewer steps and generally more art. If this is so, what are the reliability implications of such products?
- 2-18 Some polymers degrade upon exposure to radiation due to bond breaking. Covalent bond energies for some important atomic pairings are:
- | | | | |
|-----------|------------------|-----------|------------------|
| Bond pair | Energy (kJ /mol) | Bond pair | Energy (kJ /mol) |
| C-C | 340 | C-H | 420 |
| C-C | 620 | C-Cl | 330 |
- What portion of the electromagnetic spectrum would be expected to cause radiation-damage effects in each bond type?
- 2-19 Damage to opto-electronic devices and materials often stems from recombination of electron-hole pairs without photon emission (nonradiative recombination) and from self-absorption of the involved radiation. Explain how each phenomenon could cause degradation.

REFERENCES

-
1. S. M. Sze, *Semiconductor Devices—Physics and Technology*, Wiley, New York (1985).

2. B. G. Streetman, *Solid State Electronic Devices*, 4th ed., Prentice Hall, Engelwood Cliffs, NJ (1995).
3. J. M. Mayer and S. S. Lau, *Electronic Materials Science—For Integrated Circuits in Si and GaAs*, Macmillan, New York (1990).
4. A. Bar-Lev, *Semiconductors and Electronic Devices*, 3rd ed., Prentice Hall, Engelwood Cliffs, NJ (1993).
5. T. M. Fredericksen, *Intuitive IC Electronics*, 2nd ed. McGraw Hill, New York (1989).
6. S. M. Sze, ed. *VLSI Technology*, 2nd ed., McGraw Hill, New York (1988).
7. S. Middleman and A.K. Hochberg, *Process Engineering Analysis in Semiconductor Device Fabrication*, McGraw Hill, New York (1993).
8. W. S. Ruska, *Microelectronic Processing*, McGraw Hill, New York (1987).
9. G. E. Anner, *Planar Processing Primer*, Van Nostrand Reinhold, New York (1990).
10. R. C. Jaeger, *Introduction to Microelectronic Fabrication*, volume V, Addison-Wesley, Menlo Park, CA (1988).
11. A. A. Iranmanesh and B. Bastani, *IEEE Circuits and Devices* **8** (3), 14 (1992).
12. E. Adler et. al., *IBM J. Res. and Devel.* **39** (1/2), 167 (1995).
13. H. E. Mass, G. Groeseneken, H. Labon, and J. Witters, *Microelectronics Journal* **20**, 9 (1989).
14. W. T. Anderson, *Tutorial Notes, IEEE International Reliability Physics Symposium*, 7.1 (1992).
15. W. J. Roesch, *Tutorial Notes, IEEE International Reliability Physics Symposium*, 1.1 (1988).
16. A. Christou, *Tutorial Notes, IEEE International Reliability Physics Symposium*, 3.1 (1986).
17. H. Morkoc, *IEEE Circuits and Devices* **7**, 15 (1991).
18. F. Magistrali, C. Tedesco, E. Zanoni, and C. Canali, in *Reliability of Gallium Arsenide MMICs*, ed. by A. Christou, Wiley, Chichester (1992).
19. K. Wilson, *GEC J. Res.* **4**, 126 (1986).
20. A. Bensoussan, P. Coval, W. J. Roesch, and T. Rubalcava, 32nd Annual Proceedings of the IEEE Reliability Physics Symposium, p. 434 (1994).
21. C. P. Skrimshire, J. R. Farr, D. F. Sloan, M. J. Robertson, P. A. Putland, J. C. D. Stokoe, and R. R. Sutherland, *IEE Proceedings*, 137, Pt J, No 1, 74 (1990).
22. S. P. Sim, *Semiconductor Device Reliability*, ed. by A. Christou and B. A. Unger, Kluwer, Amsterdam (1990).
23. M. Fukuda, *IEEE J. Lightwave Technol.* **6**, 1488 (1988).
24. M. Fukuda, *Reliability and Degradation of Semiconductor Lasers and LEDs*, Artech, Boston (1991).

25. J. Spencer and R. Koelbl, *Tutorial Notes, IEEE International Reliability Physics Symposium*, 1.1 (1989).
26. R. W. Mann, L. A. Clevinger, P. D. Agnello, and F. R. White, *IBM J. Res. and Develop.* **39**, 403 (1995).
27. W. Zulehner in *Semiconductor Silicon*, ed. G. Harbeke and M. J. Schulz, P.2, Springer Verlag, Berlin (1989).
28. A. B. Joshi, R. Mann, L. Chung, M. Bhat, H. T. Cho, B. W. Min, and D. L. Kwong, *33rd Annual Proceedings of the IEEE Reliability Physics Symposium*, p. 156 (1995).
29. C. Felsch and E. Rosenbaum, *33rd Annual Proceedings of the IEEE Reliability Physics Symposium*, p. 142 (1995).
30. J. G. Ryan, R. M. Geffken, N. R. Poulin and J. R. Paraszcza, *IBM J. Res. and Develop.* **39**, 371 (1995).
31. J. A. Reynolds, *Solid State Tech.* **37**, No 12, 63 (1994).