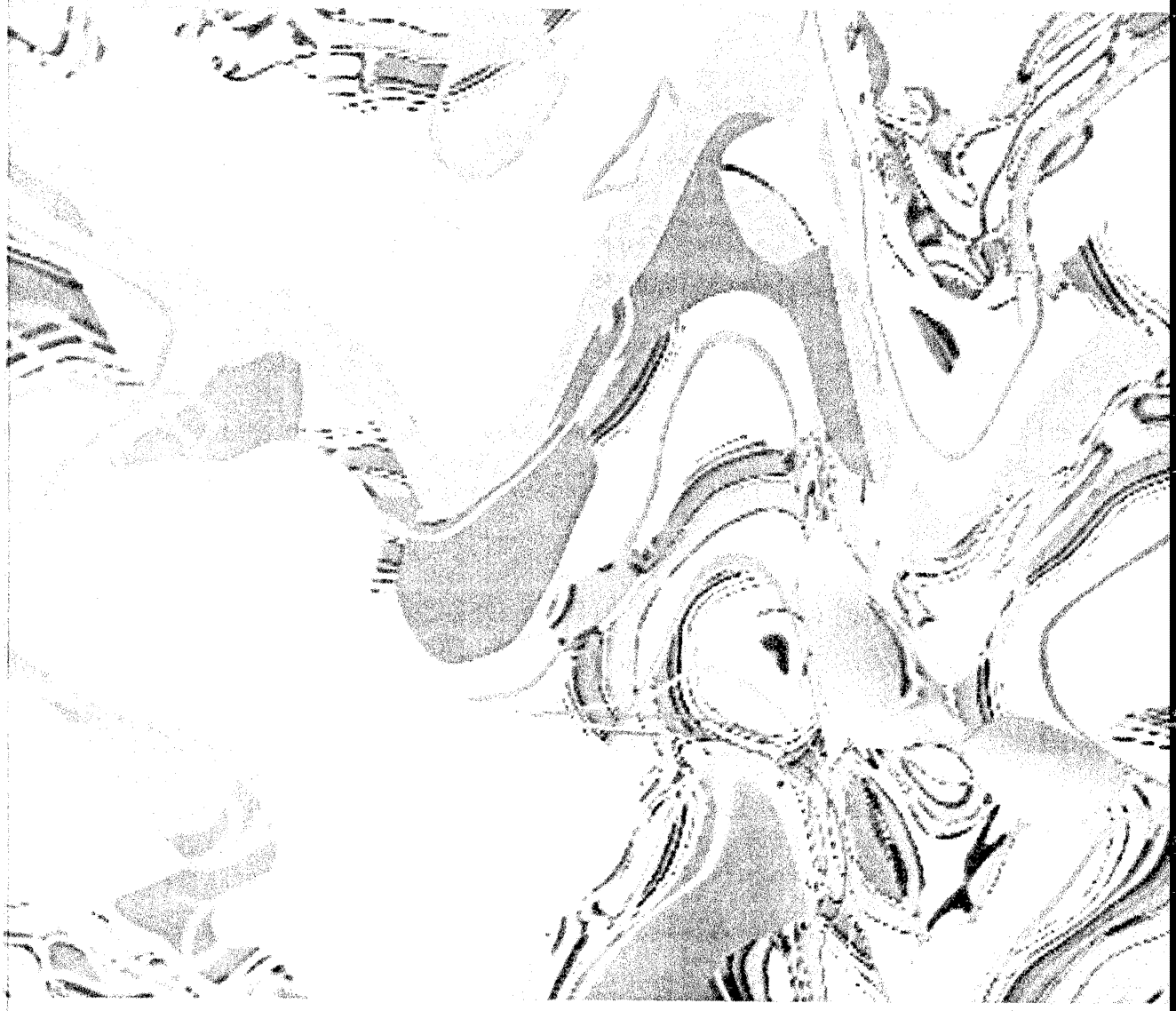
Semiconductor Manufacturing Technolo



Michael Quirk
Julian Serda

UNIVERSITY OF VIRGINIA LIBRARY
X004472523

TSMC-1014
U.S. Patent 9,337,193

For more information about this book or any of Prentice Hall's other new technology titles, visit our website at http://www.prenhall.com

Semiconductor Manufacturing Technology

This book is written for students in two- and four-year technology programs at community colleges and universities. Chapters are organized around the broad technologies applicable to semiconductor manufacturing.

Chapters 1 to 8 present fundamental technical information relevant to semiconductor manufacturing. Chapter 9 presents a process model overview with a general flowchart that links the major areas in wafer fabrication. Chapters 10 to 19 cover each of the major processes. Each process chapter concludes with a summary of quality measures and troubleshooting issues to provide the student with practical challenges encountered during wafer fabrication. Chapter 20 provides an overview of the back-end process for IC assembly and packaging.

Academic and industry reviewers have applauded this book as the most comprehensive and up-to-date text currently available in the market.

ISBN 0-13-081520-9
90000>
9 780130 815200

tice
SC
TK
78
.Q

SEMICONDUCTOR MANUFACTURING TECHNOLOGY

MICHAEL QUIRK
STAKTEK CORPORATION

JULIAN SERDA
ADVANCED MICRO DEVICES



Upper Saddle River, New Jersey
Columbus, Ohio

Library of Congress Cataloging-in-Publication Data
Quirk, Michael
Semiconductor manufacturing technology / Michael Quirk, Julian Serda.
p. cm.
ISBN 0-13-081520-9
1. Electronic industries. 2. Semiconductors—Design and construction. I. Serda, Julian.
II. Title.

TK7836.Q27 2001

00-041653

Vice President and Publisher: Dave Garza
Editor in Chief: Stephen Helba
Assistant Vice President and Publisher: Charles E. Stewart, Jr.
Production Editor: Alexandrina Benedicto Wolf
Production Coordination: Lithokraft II
Design Coordinator: Robin G. Chukes
Cover Designer: Linda Fares
Cover Image: John Foxx
Production Manager: Matthew Ottenweller
Marketing Manager: Barbara Rose

This book was set in Times Roman by Lithokraft II. It was printed and bound by R. R. Donnelley & Sons Company. The cover was printed by Phoenix Color Corp.

SCIENG
TK
7836
.Q27
2001

Quirk, Michael, 1953-
Semiconductor
manufacturing
technology

Copyright © 2001 by Prentice-Hall, Inc., Upper Saddle River, New Jersey 07458. All rights reserved. Printed in the United States of America. This publication is protected by Copyright and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permission(s), write to: Rights and Permissions Department.



10 9 8 7 6 5 4 3 2 1
ISBN 0-13-081520-9

CONTENTS

CHAPTER 1 INTRODUCTION TO THE SEMICONDUCTOR INDUSTRY

OBJECTIVES 1
INTRODUCTION 2
DEVELOPMENT OF AN INDUSTRY 2
- Industry Roots 2
- The Solid State 3

CIRCUIT INTEGRATION 4
- Integration Eras 5

IC FABRICATION 6
- Wafer Fab 7
- Stages of IC Fabrication 7

SEMICONDUCTOR TRENDS 9
- Increase in Chip Performance 9
- Increase in Chip Reliability 12
- Reduction in Chip Price 12

THE ELECTRONIC ERA 14
- The 1950s: Transistor Technology 14
- The 1960s: Process Technology 14
- The 1970s: Competition 14
- The 1980s: Automation 15
- The 1990s: Volume Production 15

CAREERS IN SEMICONDUCTOR MANUFACTURING 16
- Technician 16
- Job Descriptions 18

SUMMARY 19
KEY TERMS 19
REVIEW QUESTIONS 20
SELECTED INDUSTRY WEB SITES 20
REFERENCES 20

CHAPTER 2 CHARACTERISTICS OF SEMICONDUCTOR MATERIALS

OBJECTIVES 21
INTRODUCTION 21
ATOMIC STRUCTURE 22
- Electrons 22

THE PERIODIC TABLE 24
- Ionic Bonds 26
- Covalent Bonds 28

CLASSIFYING MATERIALS 28
- Conductors 29
- Insulators 30
- Semiconductors 32

SILICON 33
- Pure Silicon 33
- Why Silicon? 33
- Doped Silicon 34
- pn Junctions 39

ALTERNATIVE SEMICONDUCTOR MATERIALS 39
- Gallium Arsenide (GaAs) 40

SUMMARY 41
KEY TERMS 41
REVIEW QUESTIONS 42
REFERENCES 42

CHAPTER 3 DEVICE TECHNOLOGIES

OBJECTIVES 43
INTRODUCTION 43
CIRCUIT TYPES 44
- Analog Circuits 44
- Digital Circuits 44

PASSIVE COMPONENT STRUCTURES 45
- IC Resistor Structures 45
- IC Capacitor Structures 46

ACTIVE COMPONENT STRUCTURES 46
- The pn Junction Diode 46
- The Bipolar Junction Transistor 49
- Schottky Diode 51
- Bipolar IC Technology 52
- CMOS IC Technology 52

Enhancement and Depletion-Mode MOSFETs 58
LATCHUP IN CMOS DEVICES 59
INTEGRATED CIRCUIT PRODUCTS 59
Linear IC Product Types 60
Digital IC Product Types 60
SUMMARY 62
KEY TERMS 62
REVIEW QUESTIONS 63
IC MANUFACTURERS' WEB SITES 64
REFERENCES 65

CHAPTER 4 SILICON AND WAFER PREPARATION
OBJECTIVES 67
INTRODUCTION 67
SEMICONDUCTOR-GRADE SILICON 68
CRYSTAL STRUCTURE 68
Amorphous Materials 69
Unit Cells 69
Polycrystal and Monocrystal Structures 70
CRYSTAL ORIENTATION 71
MONOCRYSTAL SILICON GROWTH 72
CZ Method 72
Float-Zone Method 75
Reasons for Larger Ingot Diameters 76
CRYSTAL DEFECTS IN SILICON 78
Point Defects 78
Dislocations 79
Gross Defects 80
WAFER PREPARATION 80
Shaping Operations 81
Wafer Slicing 82
Wafer Lapping and Edge Contour 82
Etching 83
Polishing 83
Cleaning 84
Wafer Evaluation 84
Packaging 84
QUALITY MEASURES 84
Physical Dimensions 85
Flatness 86
Microroughness 86
Oxygen Content 86
Crystal Defects 87
Particles 87
Bulk Resistivity 87
EPITAXIAL LAYER 87
SUMMARY 88
KEY TERMS 88
REVIEW QUESTIONS 89
SELECTED INDUSTRY WEB SITES 90
REFERENCES 90

CHAPTER 5 CHEMICALS IN SEMICONDUCTOR FABRICATION
OBJECTIVES 91
INTRODUCTION 91
STATES OF MATTER 91
PROPERTIES OF MATERIALS 92
Chemical Properties for Semiconductor Manufacturing 93
PROCESS CHEMICALS 99
Liquids 99
Gases 103
SUMMARY 109
KEY TERMS 109
REVIEW QUESTIONS 110
CHEMICAL SUPPLIERS' WEB SITES 110
REFERENCES 111

CHAPTER 6 CONTAMINATION CONTROL IN WAFER FABS
OBJECTIVES 113
INTRODUCTION 113
Clean Background 114
TYPES OF CONTAMINATION 114
Particles 115
Metallic Impurities 116
Organic Contamination 117
Native Oxides 118
Electrostatic Discharge 119
SOURCES AND CONTROL OF CONTAMINATION 120
Air 120
Humans 121
Facility 123
Water 126
Process Chemicals 130

Production Equipment 131
Workstation Design 131
WAFER WET CLEANING 135
Wet-Cleaning Overview 135
Wet-Clean Equipment 138
Alternatives to RCA Clean 142
SUMMARY 143
KEY TERMS 143
REVIEW QUESTIONS 144
CHEMICAL AND EQUIPMENT SUPPLIERS' WEB SITES 145
REFERENCES 146

CHAPTER 7 METROLOGY AND DEFECT INSPECTION

OBJECTIVES 149
INTRODUCTION 149
IC METROLOGY 150
Measurement Equipment 151
Yield 151
Data Management 152
QUALITY MEASURES 152
Film Thickness 153
Film Stress 158
Refractive Index 158
Dopant Concentration 159
Unpatterned Surface Defects 161
Patterned Surface Defects 164
Critical Dimension (CD) 165
Step Coverage 167
Overlay Registration 167
Capacitance-Voltage (C-V) Test 168
Contact Angle 171
ANALYTICAL EQUIPMENT 171
Secondary-Ion Mass Spectrometry (SIMS) 171
Atomic Force Microscope (AFM) 173
Auger Electron Spectroscopy (AES) 174
X-Ray Photoelectron Spectroscopy (XPS) 174
Transmission Electron Microscope (TEM) 175
Energy- and Wavelength-Dispersive Spectrometer (EDX and WDX) 176
Focused Ion Beam (FIB) 176
SUMMARY 177
KEY TERMS 178
REVIEW QUESTIONS 178
METROLOGY EQUIPMENT SUPPLIERS' WEB SITES 179
REFERENCES 180

CHAPTER 8 GAS CONTROL IN PROCESS CHAMBERS

OBJECTIVES 181
INTRODUCTION 181
VACUUM 183
Vacuum Ranges 183
Mean Free Path 184
VACUUM PUMPS 184
Roughing Pump 185
High Vacuum Pump 186
Vacuum in Integrated Tools 188
PROCESS CHAMBER GAS FLOW 189
Mass Flow Controllers 189
RESIDUAL GAS ANALYZER (RGA) 190
RGA Basics 190
RGA as Real-Time Monitor 191
PLASMA 192
Glow Discharge 193
PROCESS CHAMBER CONTAMINATION 195
SUMMARY 196
KEY TERMS 196
REVIEW QUESTIONS 196
VACUUM EQUIPMENT SUPPLIERS' WEB SITES 197
REFERENCES 197

CHAPTER 9 IC FABRICATION PROCESS OVERVIEW

OBJECTIVES 199
INTRODUCTION 199
CMOS PROCESS FLOW 199
Overview of Areas in a Wafer Fab 200
CMOS MANUFACTURING STEPS 205
1. Twin Well Process 205
2. Shallow Trench Isolation Process 207
3. Poly Gate Structural Process 210
4. Lightly Doped Drain (LDD) Implant Process 211
5. Sidewall Spacer Formation 212
6. Source/Drain (S/D) Implant Processes 213

7. Contact Formation 214
8. Local Interconnect (LI) Process 214
9. Via-1 and Plug-1 Formation 216
10. Metal-1 Interconnect Formation 217
11. Via-2 and Plug-1 Formation 218
12. Metal-2 Interconnect Formation 219
13. Metal-3 to Pad Etch and Alloy 220
14. Parametric Testing 221

SUMMARY 222
KEY TERMS 223
REVIEW QUESTIONS 223
REFERENCES 224

CHAPTER 10 OXIDATION

OBJECTIVES 225
INTRODUCTION 225
OXIDE FILM 226
Nature of Oxide Film 226
Uses of Oxide Film 227
THERMAL OXIDATION GROWTH 231
Chemical Reaction for Oxidation 231
Oxidation Growth Model 232
FURNACE EQUIPMENT 239
HORIZONTAL VERSUS VERTICAL FURNACES 240
Vertical Furnace 241
Fast Ramp Vertical Furnace 245
Rapid Thermal Processor 246
OXIDATION PROCESS 248
Pre Oxidation Cleaning 248
Oxidation Process Recipe 249
QUALITY MEASUREMENTS 250
OXIDATION TROUBLESHOOTING 252
SUMMARY 252
KEY TERMS 253
REVIEW QUESTIONS 253
FURNACE AND RTP EQUIPMENT SUPPLIERS' WEB SITES 254
REFERENCES 255

CHAPTER 11 DEPOSITION

OBJECTIVES 257
INTRODUCTION 257
Film Layering Terminology 258
FILM DEPOSITION 260
Thin-Film Characteristics 260
Film Growth 263
Film Deposition Techniques 264
CHEMICAL VAPOR DEPOSITION 265
CVD Chemical Processes 265
CVD Reaction 266
CVD DEPOSITION SYSTEMS 269
CVD Equipment Design 270
APCVD (Atmospheric Pressure CVD) 271
LPCVD (Low Pressure CVD) 273
Plasma-Assisted CVD 277
DIELECTRICS AND PERFORMANCE 282
Dielectric Constant 282
Device Isolation 286
SPIN-ON-DIELECTRICS 287
Spin-On-Glass (SOG) 287
Spin-On-Dielectric (SOD) 288
EPITAXY 289
Epitaxy Growth Methods 290
CVD QUALITY MEASURES 292
CVD TROUBLESHOOTING 292
SUMMARY 294
KEY TERMS 295
REVIEW QUESTIONS 295
DEPOSITION EQUIPMENT SUPPLIERS' WEB SITES 296
REFERENCES 296

CHAPTER 12 METALLIZATION

OBJECTIVES 299
INTRODUCTION 300
TYPES OF METALS 302
Aluminum 302
Aluminum-Copper Alloys 305
Copper 305
Barrier Metals 307
Silicides 309
Metal Plugs 312
METAL DEPOSITION SYSTEMS 313
Evaporation 313
Sputtering 314
Metal CVD 320
Copper Electroplate 323
METALLIZATION SCHEMES 325
Traditional Aluminum Structure 325

Copper Damascene Structure 326
METALLIZATION QUALITY MEASURES 329
METALLIZATION TROUBLESHOOTING 330
SUMMARY 331
KEY TERMS 331
REVIEW QUESTIONS 332
METALLIZATION EQUIPMENT AND MATERIALS SUPPLIERS' WEB SITES 332
REFERENCES 333

CHAPTER 13 PHOTOLITHOGRAPHY: VAPOR PRIME TO SOFT BAKE

OBJECTIVES 335
INTRODUCTION 336
Photolithography Concepts 336
PHOTOLITHOGRAPHY PROCESSES 339
Negative Lithography 339
Positive Lithography 340
EIGHT BASIC STEPS OF PHOTOLITHOGRAPHY 342
Step 1: Vapor Prime 343
Step 2: Spin Coat 343
Step 3: Soft Bake 344
Step 4: Alignment and Exposure 344
Step 5: Post-Exposure Bake (PEB) 344
Step 6: Develop 344
Step 7: Hard Bake 344
Step 8: Develop Inspect 345
VAPOR PRIME 345
Wafer Cleaning 345
Dehydration Bake 346
Wafer Priming 346
SPIN COAT 348
Photoresist 348
Photoresist Physical Properties 349
Conventional I-Line Photoresists 351
Deep UV (DUV) Photoresists 354
Photoresist Dispensing Methods 357
SOFT BAKE 360
Soft Bake Equipment 360
Process Characterization 361
PHOTORESIST QUALITY MEASURES 362
PHOTORESIST TROUBLESHOOTING 363
SUMMARY 364
KEY TERMS 364
REVIEW QUESTIONS 364
PHOTORESIST MATERIALS AND EQUIPMENT SUPPLIERS' WEB SITES 365
REFERENCES 366

CHAPTER 14 PHOTOLITHOGRAPHY: ALIGNMENT AND EXPOSURE

OBJECTIVES 367
INTRODUCTION 367
Importance of Alignment and Exposure 368
OPTICAL LITHOGRAPHY 370
Light 370
Exposure Sources 372
Optics 376
Resolution 385
PHOTOLITHOGRAPHY EQUIPMENT 388
Contact Aligner 388
Proximity Aligner 389
Scanning Projection Aligner 390
Step-and-Repeat Aligner (Stepper) 390
Step-and-Scan System 393
Reticles 394
Optical Enhancement Techniques 398
Alignment 400
Environmental Conditions 404
Comparison of Photo Tools 405
MIX AND MATCH 406
ALIGNMENT AND EXPOSURE QUALITY MEASURES 407
ALIGNMENT AND EXPOSURE TROUBLESHOOTING 408
SUMMARY 409
KEY TERMS 410
REVIEW QUESTIONS 410
PHOTORESIST MATERIALS AND EQUIPMENT SUPPLIERS' WEB SITES 411
REFERENCES 412

CHAPTER 15 PHOTOLITHOGRAPHY: PHOTORESIST DEVELOPMENT AND ADVANCED LITHOGRAPHY

OBJECTIVES 413
INTRODUCTION 413
Advanced Lithography 414
POST-EXPOSURE BAKE 414
DUV Post-Exposure Bake (PEB) 415
Conventional I-Line PEB 416
DEVELOP 416
Negative Resist 417
Positive Resist 418
Development Methods 419
Resist Development Parameters 421
HARD BAKE 422
DEVELOP INSPECT 422
ADVANCED LITHOGRAPHY 424
Next-Generation Lithography 424
Advanced Resist Processing 428
DEVELOP QUALITY MEASURES 429
DEVELOP TROUBLESHOOTING 431
SUMMARY 432
KEY TERMS 433
REVIEW QUESTIONS 433
PHOTOLITHOGRAPHY MATERIALS AND EQUIPMENT SUPPLIERS' WEB SITES 433
REFERENCES 434

CHAPTER 16 ETCH

OBJECTIVES 435
INTRODUCTION 435
Etch Processes 436
ETCH PARAMETERS 437
Etch Rate 437
Etch Profile 438
Etch Bias 439
Selectivity 440
Uniformity 441
Residues 441
Polymer Formation 442
Plasma-Induced Damage 442
Particle Contamination 443
DRY ETCH 443
Etching Action 444
Potential Distribution 444
PLASMA ETCH REACTORS 446
Barrel Plasma Etcher 446
Parallel Plate (Planar) Reactor 447
Downstream Etch Systems 448
Triode Planar Reactor 448
Ion Beam Milling 448
Reactive Ion Etch (RIE) 450
High-Density Plasma Etchers 450
Etch System Review 453
Endpoint Detection 455
Vacuum for Etch Chambers 456
DRY ETCH APPLICATIONS 456
Dielectric Dry Etch 457
Silicon Dry Etch 459
Metal Dry Etch 462
WET ETCH 464
Types of Wet Etch 465
HISTORICAL PERSPECTIVE 466
PHOTORESIST REMOVAL 466
Plasma Ashing 466
ETCH INSPECTION 469
ETCH INSPECTION QUALITY MEASURES 469
DRY ETCH TROUBLESHOOTING 470
SUMMARY 471
KEY TERMS 472
REVIEW QUESTIONS 472
ETCH EQUIPMENT SUPPLIERS' WEB SITES 473
REFERENCES 474

CHAPTER 17 ION IMPLANT

OBJECTIVES 475
INTRODUCTION 475
Doped Regions 477
DIFFUSION 479
Diffusion Principles 479
Diffusion Process 481
ION IMPLANTATION 482
Overview 483
Ion Implant Parameters 484
ION IMPLANTERS 488
Ion Source 488
Extraction and Ion Analyzer 490
Acceleration Column 491

Scanning System 494
Process Chamber 498
Annealing 499
Channeling 501
Particles 502
ION IMPLANT TRENDS IN PROCESS INTEGRATION 503
Deep Buried Layers 503
Retrograde Wells 503
Punchthrough Stoppers 504
Threshold Voltage Adjustment 504
Lightly Doped Drain 505
Source/Drain Implants 505
Polysilicon Gate 506
Trench Capacitor 506
Ultrashallow Junctions 506
Silicon-On-Insulator (SOI) 507
ION IMPLANT QUALITY MEASURES 508
ION IMPLANT TROUBLESHOOTING 509
SUMMARY 510
KEY TERMS 511
REVIEW QUESTIONS 512
ION IMPLANTER EQUIPMENT SUPPLIERS' WEB SITES 513
REFERENCES 513

CHAPTER 18 CHEMICAL MECHANICAL PLANARIZATION

OBJECTIVES 515
INTRODUCTION 515
TRADITIONAL PLANARIZATION 518
Etchback 518
Glass Reflow 519
Spin-On Films 520
CHEMICAL MECHANICAL PLANARIZATION 520
CMP Planarity 521
Advantages of CMP 522
CMP Mechanisms 523
CMP Slurry and Pad 526
CMP Equipment 529
CMP Clean 532
CMP Equipment Manufacturers 534
CMP APPLICATIONS 535
STI Oxide Polish 535
LI Oxide Polish 535
LI Tungsten Polish 536
ILD Oxide Polish 536
Tungsten Plug Polish 536
Dual-Damascene Copper Polish 537
CMP QUALITY MEASURES 538
CMP TROUBLESHOOTING 540
SUMMARY 541
KEY TERMS 541
REVIEW QUESTIONS 542
CMP EQUIPMENT SUPPLIERS' WEB SITES 542
REFERENCES 543

CHAPTER 19 WAFER TEST

OBJECTIVES 545
INTRODUCTION 545
IC Electrical Tests 546
WAFER TEST 547
In-Line Parametric Test 547
Wafer Sort 555
Yield 560
Wafer Sort Yield Models 563
TEST QUALITY MEASURES 565
TEST TROUBLESHOOTING 567
SUMMARY 567
KEY TERMS 568
REVIEW QUESTIONS 568
TEST AND PROBER EQUIPMENT SUPPLIERS' WEB SITES 569
REFERENCES 569

CHAPTER 20 ASSEMBLY AND PACKAGING

OBJECTIVES 571
INTRODUCTION 571
Packaging Levels 573
TRADITIONAL ASSEMBLY 574
Backgrind 574
Die Separation 575
Die Attach 575
Wirebonding 577
TRADITIONAL PACKAGING 580
Plastic Packaging 581
Ceramic Packaging 584
Final Test 586

ADVANCED ASSEMBLY AND PACKAGING 586
- Flip Chip 587
- Ball Grid Array (BGA) 589
- Chip on Board (COB) 590
- Tape Automated Bonding (TAB) 591
- Multichip Module (MCM) 591
- Chip Scale Packaging (CSP) 592
- Wafer-Level Packaging 592

ASSEMBLY AND PACKAGING QUALITY MEASURES 596

IC PACKAGING TROUBLESHOOTING 597

SUMMARY 597

KEY TERMS 598

REVIEW QUESTIONS 598

ASSEMBLY AND PACKAGING SUPPLIERS' WEB SITES 599

REFERENCES 600

APPENDICES

APPENDIX A CHEMICALS AND SAFETY 601

APPENDIX B CONTAMINATION CONTROLS IN CLEANROOMS 610

APPENDIX C UNITS 614

APPENDIX D COLOR AS A FUNCTION OF OXIDE THICKNESS 617

APPENDIX E OVERVIEW OF PHOTORESIST CHEMISTRY 618

APPENDIX F ETCH CHEMISTRY 622

GLOSSARY 624

INDEX 648

PREFACE

This text started with a simple premise: as instructors, we need to teach relevant microchip technology to students and employees in semiconductor manufacturing. Unfortunately, in the semiconductor industry, changes in technology are measured in months, not years. Our challenge was to write a relevant book that would not be outdated by the time it was published. With that in mind, we researched the material and applied ourselves to writing the chapters and creating the artwork. Following the aggressive pace of Moore's law, the technical material in our book is at most only 18 to 24 months old. This permits us to keep abreast of the changing technology nodes swirling through the semiconductor industry.

This text is written for students in two-year and four-year technology programs at community colleges and universities. The text will also be a practical reference as well as a standard text in corporate and technical training classes. Students are expected to have an understanding of high school chemistry, physics, and math. Chapters are organized around the broad technologies applicable to semiconductor manufacturing.

ORGANIZATION OF THE TEXT

Our goal is to accomplish three objectives:

1. Help technology students grasp the fundamental technologies used in manufacturing semiconductor devices.
2. Present some of the many challenges in microchip fabrication.
3. Instill in the reader an appreciation of the conceptual simplicity of semiconductor manufacturing.

All fundamental technical information relevant to semiconductor manufacturing is first presented in Chapters 1 to 8. Chapter 9 presents a process model overview with a general flowchart that links the major areas in a wafer fab. Chapters 10 to 19 cover each of the major processes in the fab. Finally, Chapter 20 provides an overview of the back-end process for IC assembly and packaging. The content in the process chapters (Chapters 10 to 20) addresses critical process technology, followed by the various equipment designs needed to support this technology. Each process chapter concludes with a summary of quality measures and troubleshooting issues to familiarize the student with the practical, day-to-day challenges encountered during wafer fabrication.

The latest technologies for sub-0.25 μm processing are covered in detail. This includes chemical mechanical planarization (CMP), shallow trench isolation (STI), chemically amplified deep UV photoresists, step-and-scan systems, copper metallization with dual damascene, and the widespread move to process integration with cluster tools. Throughout the text, we explain all process and equipment technology in light of the long history of change in the industry. Early tools and processes are described to clarify the development of current technology. In some cases, the linkage between the latest equipment and earlier tools is obvious, while in other instances the change is dramatic.

Professors, students, and other readers of this book can send comments or questions about this text to the authors at the following website: http://www.smtbook.com. We look forward to any exchange of information that can help advance semiconductor manufacturing education.

ACKNOWLEDGMENTS

We would like to thank the following reviewers for their valuable feedback: Bruce Bothwell, Chemeketa Community College; Mike Delia, Samsung Austin Semiconductor; Chris Dennis and David Hata, Portland Community College; John Fowler, Arizona State University; Lance Kinney; Bassam Matar, Maricopa Community College; John Nistler, Advanced Micro Devices; Val Shires, Gwinnett Technical Institute; and Joseph White, Texas A&M University.

We acknowledge the support of the following colleagues, friends, and new acquaintances for their technical inputs and encouragement throughout the project.

Technical Reviewers and Editors:

- Mike D'Elia, Member of Technical Staff, Advanced Micro Devices—a tremendous inspiration to us—reviewed the process flow, oxidation and deposition chapters and was the first SMT instructor to use selected chapters.
- Mike Hillis, Module Shift Manager, Advanced Micro Devices, provided encouragement and reviewed and edited the photolithography sections.
- Steve Hymes, Ph.D., Applications Manager, Ashland Specialty Chemicals, who seemed more excited than us at times, reviewed and edited the CMP and metallization chapters.
- Rick Jarvis, Senior Member Technical Staff, Advanced Micro Devices, on assignment to International SEMATECH, with great insight to the industry, reviewed and edited the contamination, etch, and deposition sections.
- Larry Knipp, President, Nykar Inc., reviewed and edited the RF and plasma sections of the book.
- Jeremy Lansford, Module Manager, Advanced Micro Devices, very carefully reviewed and edited the CMP section, then met with us to ensure his suggestions were understood.
- Federico Miller, Chemistry Department, Austin Community College, provided advice and encouragement throughout the project and reviewed all the chemistry related information.
- Sergei Postnikov, Ph.D., Photolithography Development, Motorola, reviewed the photolithography sections and provided encouragement throughout the project.
- Alan Romriell, Senior Test Engineer, Advanced Micro Devices, reviewed the test chapter.
- Andy Shurtleff, President, Liquid Asset Services, an expert in ultra-pure water systems, reviewed and edited the section on ultra-pure water.
- Darby Webster, Senior Manager, Manufacturing Engineering, Staktek Corp., reviewed the assembly and packaging chapter and gave his support throughout the project.
- Joseph W. Wiseman, Member of Technical Staff, Advanced Micro Devices, a professional in his engineering approach and also an etch instructor, reviewed and edited the etch section.

Computer-Aided Illustrators and Writers:

- Thuat Duong, CAD specialist, Solectron, created many of the drawings used in the assembly and packaging section.
- Matt Hanson, Computer Graphics Artist, provided graphic support for the metrology section.
- Josh Hix, Technical Writer, a full-time engineering student at Georgia Tech University, produced most of the glossary.
- Trent Romriell, Computer Graphics Artist, supported us with many drawings in the beginning chapters.
- James Samara, Computer Graphics Artist, provided graphics support for the metrology chapter.
- Mark Serda, Computer Graphics Artist, provided graphics support for the gas controls chapter.

Providers of Logistics, Technical Information, Graphics, and Photographs:

- Shannon Bacon, Continuous Improvement Facilitator, Advanced Micro Devices, made arrangements for us to meet with subject matter experts at AMD's facilities.
- Tony Denboer, Integrated Circuit Engineering Corporation, provided SEM micrographs of actual integrated circuit structures.
- Jim DeVries, Facilities Operations Manager, Advanced Micro Devices, directed the photo shoot of some AMD facilities.
- Charlie Earnhart, Vice-President, Staktek Corp., for his support throughout the project.
- James A. Elliott, Technical Instructor, Varian Semiconductor Equipment, assisted the writers with obtaining permission to use technical documentation and illustrations of Varian ion implanters.
- Robert Emmons, Learning and Development Manager, Advanced Micro Devices, allowed Julian to adjust his work hours so he could devote more quality time to the book.
- Kathleen Fitts, Global Quality Programs Manager, Applied Materials, supported the project with various diagrams and photographs.
- Dahlia Hernandez-Fridge, Executive Assistant to the President, CEDRA Corp., researched technical information.
- Van Pham, Process Development Engineer, Staktek Corporation, provided input regarding etch.
- Richard Rogoff, Major Accounts Manager, ASML, provided excellent diagrams and photographs of step-and-scan systems.
- Tim Roy, Director, Staktek Corp., for his support throughout the project.
- Sharon Shaw, Librarian, Advanced Micro Devices, showed us what the word "resourceful" means by always being able to locate a reference for us and for handling the logistics.
- Michael Sullivan, MEGASYSTEM Manager, Air Products, allowed us to photograph several bulk gas distribution sites.
- Chuck Tully, Austin Technical Training Center Manager, Applied Materials, who was always there when we needed him, provided technical information and graphics on electrostatic chuck and DPS etch equipment.
- Phil Ware, Director and General Manager of Marketing, Semiconductor Equipment Division, Canon USA, Inc., provided contacts, graphics, and photographs of several generations of photolithography tools.
- Scott A. Williams, Central Regional Account Manager, SpeedFam-IPEC, arranged to have photographs taken and provided diagrams of CMP equipment.

Michael Quirk
Julian Serda

This text is dedicated to our families, who endured so much during this project:

Monique, Justin, and Morgán Quirk

Dorothy, Mark, Ruben, and Elena Serda.

CHAPTER 1

INTRODUCTION TO THE SEMICONDUCTOR INDUSTRY

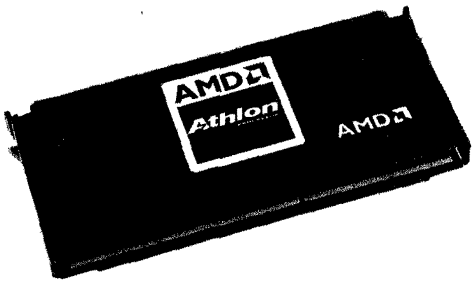
Society witnessed a technology revolution during the twentieth century with the change from products derived from mechanical technology to those centered on electronic technology. Digital compact disc players replaced record players, and automotive engines are now controlled by electronic ignition systems. Electronic computers quickly perform tasks in nearly every aspect of society, promoting efficient use of our resources. Given the breadth of changes brought by electronic technology, the revolution has just begun.

The semiconductor industry has been at the center of this technology revolution. The principal building material, the *semiconductor*, is the main ingredient of the electronic products found throughout society. Semiconductor products are manufactured by people with diverse technical skills: designers who create new designs based on customer needs, engineers who specify improved requirements for the equipment and process, and technicians who fabricate the semiconductor products in automated factories. All the while, the growing semiconductor market continues to demand more performance at lower cost.

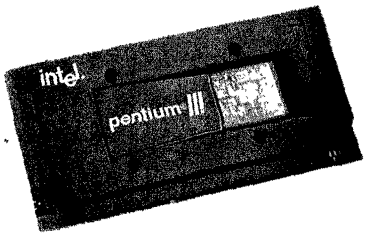
OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Describe the current economic state and the technical roots of the semiconductor industry.
2. Explain what an integrated circuit (IC) is and list the five circuit integration eras.
3. Describe a wafer, including how it is layered, and describe the essential aspects of the five stages of wafer fabrication.
4. State and discuss the three major trends associated with improvement in wafer fabrication.
5. Explain what a critical dimension (CD) is and how Moore's law predicts future wafer fabrication improvement.
6. Describe the different eras of electronics since the invention of the transistor up to modern wafer fabrication.
7. Discuss different career paths in the semiconductor industry.



Microprocessor Chip
(Photo courtesy of Advanced Micro Devices)



Microprocessor Chip
(Photo courtesy of Intel Corporation)

INTRODUCTION

The basic semiconductor material from which electronic devices are made comes in the form of round thin crystalline disks called *wafers*. Semiconductor products are produced from wafers in wafer fabrication (*wafer fab*) factories, and are referred to as *microchips*, or *chips*.

The technology of semiconductor fabrication is complex, requiring many specialized process steps, materials, equipment, and supplier industries. Once the microchips are fabricated, they are packaged into the various electronic and mechanical assemblies required for the numerous product applications. Examples of such applications are automotive electronics, electronic commerce, personal computers, and mobile cellular communications.

Worldwide sales of microchips are expected to exceed $200 billion in 2001.[1] Today, semiconductors account for 30% to 40% of the cost for a personal computer, and about $100 worth of semiconductors are found in each cellular telephone. Every automobile has approximately $140 worth of microchips, an amount that is increasing as cars become "smarter."

The semiconductor industry is actually a subset of a larger entity—the high-tech industry. Fabricating the microchips produces electronic hardware that is coupled with software to integrate and control the chip functions. The high-tech industry encompasses all the hardware and software technology found in all semiconductor applications (See Figure 1.1).

The high-tech industry in the United States is big. In the mid-1990s, the high-tech industry contributed 27% of the U.S. economy, compared to 14% for residential housing and 4% for the automotive industry. Why has the semiconductor industry become so large? A major factor has been the industry's ability to consistently increase semiconductor product performance while decreasing the price. The ability to meet the market demands for high performance at low cost can be directly traced to the technology improvements regularly made in product design and manufacturing throughout the history of the industry.

DEVELOPMENT OF AN INDUSTRY

The technology used to fabricate semiconductor devices developed out of many different technological inventions from early pioneers in the electronics field.

Industry Roots

The developmental roots of the semiconductor industry grew out of technologies developed in the first half of the twentieth century.[2] Key technical knowledge was gained from a web of industry and academia: vacuum tube electronics, wireless communications, mechanical tabulators, and solid-state physics. These industries were the fabric of the high-tech industry at the time.

The triode vacuum tube for amplifying electrical signals was invented by Lee De Forest in 1906. It was developed out of earlier vacuum tube work by John Fleming and Thomas Edison. A triode consists of three elements: two electrodes and a grid separated in an evacuated glass enclosure. A vacuum is necessary to keep the elements from burning up and also to facilitate the transfer of electrons between the electrodes. De Forest patented and named his vacuum tube invention the audion because he thought there was some potential for amplifying and reproducing sound. He

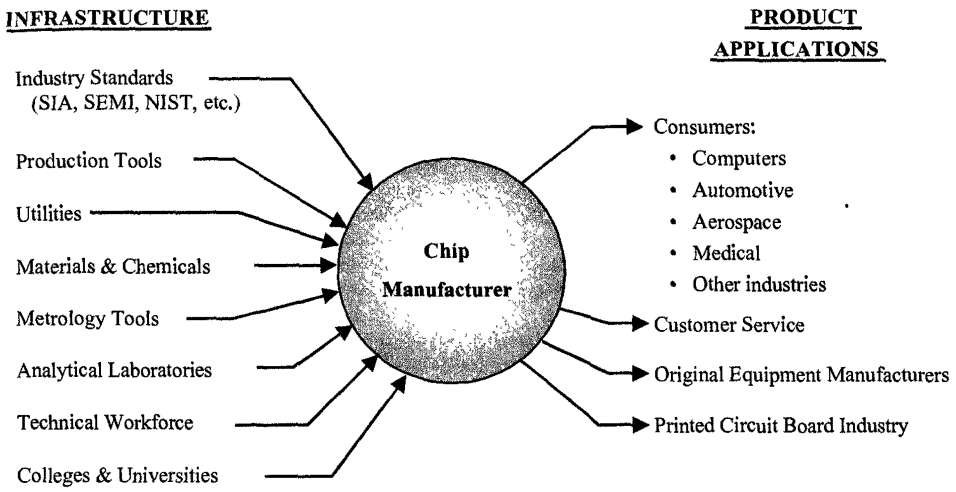


FIGURE 1.1 The Semiconductor Industry

was right, and the vacuum tube became the primary electronic device for the growth of modern radio, television, and general electronics up until the 1950s.

A relatively new material, a crystal known as silicon, was used in the early 1900s to convert wireless communications signals from alternating current to direct current. The term "semiconductor" was first adopted in Germany to cover this class of material.[3] However, to truly develop the power of semiconductor technology required additional research by physical chemists and physicists around the world. This research was needed in order to understand the quantum theory of electron behavior before a full explanation of semiconductor properties was available. This substantial work extended over several decades and beyond World War II.

There was a market need to electronically calculate numbers. The first mechanical computer was the Hollerith mechanical tabulating machine, invented by Herman Hollerith, to tabulate the 1890 census figures in an amazing six weeks. This tabulator was driven by an electric motor and based on punched cards (punch cards remained the common input method for computers up through the 1970s). Vacuum tubes were used to develop the first electronic computer, the ENIAC (Electronic Numeric Integrator and Calculator) at the University of Pennsylvania during World War II. The ENIAC weighed 50 tons, required 3,000 ft^2 of floor space, needed 19,000 vacuum tubes, and used the equivalent electrical power of 160 lighthouses.

A major drawback to the ENIAC, besides its bulky size, was the number of problems associated with vacuum tubes. Tubes were large, unreliable and consumed large amounts of electric power. Tubes had a limited lifetime due to burnout. Vacuum tubes clearly were not the optimum technology to produce small, reliable electronic products that were needed for the rapidly developing electronic markets.

The Solid State

There was a concerted effort at Bell Telephone Laboratories after World War II to study the properties of solid silicon and germanium semiconductor crystals. The scientists conducting the research sensed the need to replace vacuum tubes and the benefits from using a solid semiconductor material in place of vacuum tubes.

The modern-day semiconductor industry was born with the invention of the *solid-state transistor* at Bell Telephone Laboratories on December 16, 1947, by William Shockley, John Bardeen, and Walter Brattain.[4] The transistor, a name taken from the two terms *transconductance* and *varistor*,[5] offered the same electrical functions as a vacuum tube, but with the distinct advantages of a solid state: miniscule size, no vacuum, reliable, lightweight, minimal heat production, and low power consumption. The three scientists were awarded the 1956 Nobel Prize in physics for their invention. This discovery launched the modern semiconductor industry based on solid-state materials and technology.

The semiconductor industry started growing rapidly in the 1950s to commercialize transistor technology based on silicon. Many of the early pioneers started in northern California, in the area



Vacuum Tubes

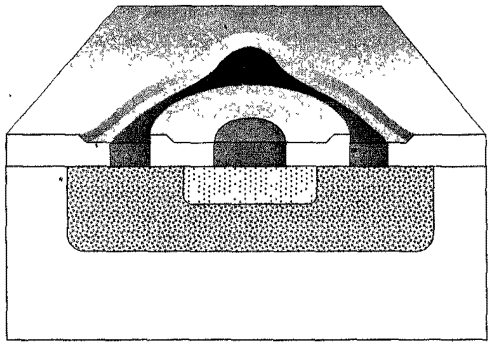


FIGURE 1.2 The First Planar Transistor

now known as Silicon Valley. The first commercial planar transistor was fabricated at Fairchild Semiconductor in Palo Alto, California, in 1957. It had a layer of aluminum interconnect material that was deposited on top of the silicon wafer to connect the different parts of the transistor (see Figure 1.2). A natural oxide layer thermally grown from the silicon was used to insulate the aluminum conductor. The use of layers was an important development in semiconductors and is the reason for the term *planar technology.*

CIRCUIT INTEGRATION

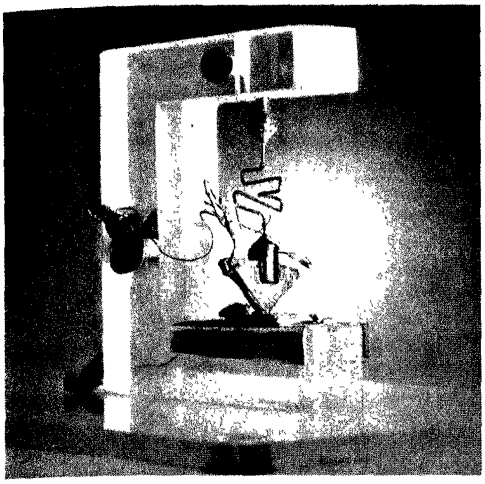
The semiconductor name derives from the material's ability to sometimes be a conductor to electricity and other times act as a nonconductor to electricity. The earliest semiconductor material was germanium, built into a single chip with one function (referred to as a discrete component). Nowadays, greater than 85% of all microchips are made from silicon semiconductor material. For this reason, we will emphasize silicon in this book.

An important step forward in the semiconductor industry was the integration of multiple electronic components on one silicon substrate. Referred to as an *integrated circuit,* or *IC,* it was coinvented independently by Robert Noyce at Fairchild Semiconductor and Jack Kilby at Texas Instruments in 1959. On the silicon surface of an IC are manufactured the many different semiconductor *devices,* such as transistors, diodes, resistors, and capacitors, which are connected into a circuit to define how the chip will function. The expression *IC* is often meant to describe the chip and all its components.

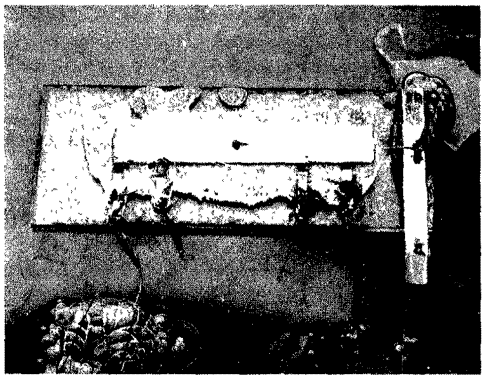
The first IC, developed in July 1958 by Jack Kilby at a Texas Instruments facility in Dallas, Texas, was made using a slice of germanium semiconductor material as the substrate.[6] The invention combined a transistor and other components on the germanium while also using the natural resistance of the germanium as a resistor. The devices were connected with individual wires.

Robert Noyce at Fairchild Semiconductor also invented the concept of an IC by extending the idea of how to interconnect the different components on the planar silicon material. His idea was to use an aluminum metal conductor on the silicon surface to interconnect the different transistors while using a layer of oxide grown from the silicon as an insulator to separate the metal conductor from the silicon devices. This was the first practical structure of an IC as a single-structure silicon chip.[7]

Circuit integration has increased dramatically since the initial IC. Because all the components are integrated on the silicon substrate, the IC has developed into a cost-effective and reliable way to produce and interconnect many components. The ability to integrate many different components on an IC has spurred engineers to design ever more complex electronic circuits to meet new customer needs.



The First Transistor from Bell Labs
(Photo courtesy of Lucent Technologies, Bell Labs Innovations)



Jack Kilby's First Integrated Circuit
(Photo courtesy of Texas Instruments, Inc.)

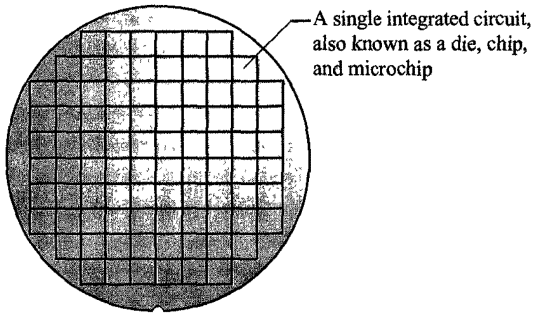


FIGURE 1.3 Top View of Wafer with Chips

Integration Eras

We can roughly categorize the integration era by the number of components integrated on a chip. This is a useful method of organizing the development of the semiconductor industry since 1960 to the present (see Table 1.1).

TABLE 1.1 Circuit Integration of Semiconductors

Circuit Integration	Semiconductor Industry Time Period	Number of Components per Chip
No integration (discrete components)	Prior to 1960	1
Small scale integration (SSI)	Early 1960s	2 to 50
Medium scale integration (MSI)	1960s to Early 1970s	50 to 5,000
Large scale integration (LSI)	Early 1970s to Late 1970s	5,000 to 100,000
Very large scale integration (VLSI)	Late 1970s to Late 1980s	100,000 to 1,000,000
Ultra large scale integration (ULSI)	1990s to present	> 1,000,000

Circuit integration continues today as the number of devices on a chip is continually increased. A significant challenge for circuit integration is the capability of the semiconductor manufacturing process to improve the fabrication technology to produce highly integrated ULSI chips at acceptable costs. To this end, the semiconductor industry has become highly standardized, with most manufacturers using similar manufacturing process and equipment technology. The key to market success is the ability of a company to deliver the right product at the right time. Industry

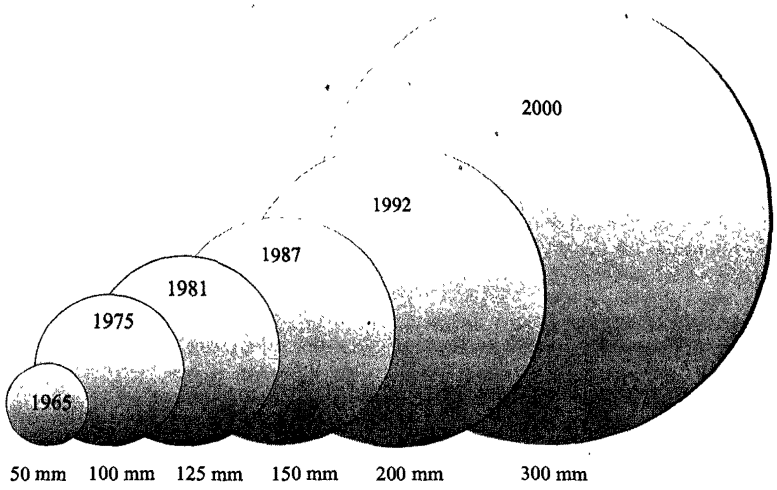


FIGURE 1.4 Evolution of Wafer Size

consortiums, such as International SEMATECH and the Semiconductor Industry Association (SIA), have been created to assist chip producers to be competitive in the world market of high-performance ICs.

IC FABRICATION

Tens or hundreds of identical chips are fabricated simultaneously on a silicon *wafer* (see Figure 1.3 on page 5). The number of chips on a wafer varies depending on the type of product and the size of each chip. The chip size changes depending on the level of integration on the chip.

The chip is also referred to as a *die* (singular and plural for chips or ICs), while the silicon wafer is often called the *substrate*. The wafer diameter has evolved over the years from an initial diameter of less than $1''^{8}$ to the present common diameter of 8″ (200 mm), which is undergoing a change to 12″, or 300 mm (see Figure 1.4). The cost of fabricating ICs drops dramatically if there are more chips on a wafer due to benefits from the economies of scale (more chips are produced from the same effort).

Fabrication of the semiconductor devices occurs only in the first few microns of the silicon near the wafer surface. The bulk of the silicon wafer thickness is to give the wafer adequate rigidity during processing. Once the devices are fabricated in the silicon, then layers of metal circuitry are placed on the silicon to interconnect between the devices and the various electronic signals outside the chip (see Figure 1.5). The concept and materials for interconnecting a modern day IC are remarkably similar to the initial planar transistor first commercialized at Fairchild Semiconductor in 1957. The major difference is that today's chips are much more complex.

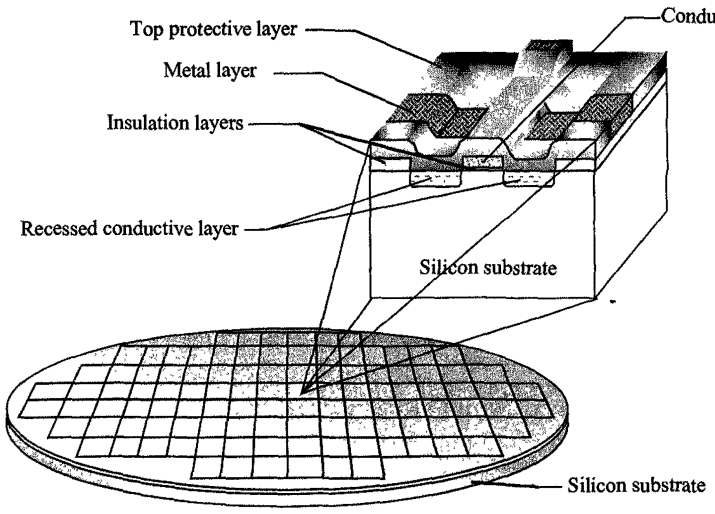


FIGURE 1.5 Devices and Layers from a Silicon Chip

Wafer Fab

Early wafer fabs were simple, with operators manually processing wafers through the different operations. An essential aspect of wafer fabs is that as wafers become more densely integrated, the permissible level of contamination decreases dramatically. Contamination that can damage wafers and cause them not to function properly evolves from many sources: humans, materials, water, air, and equipment. Modern wafer fabs have become specialized facilities that provide the clean manufacturing environment and specialized equipment to produce wafers with minimal defects from contamination. This includes limited human exposure, ultrapure chemicals and utilities, and special wafer handling procedures needed to fabricate ICs in the ULSI era.

A wafer typically requires two to three months of processing to complete the 450 or more process steps in a wafer fab. At the end of the fabrication process, the individual chips are separated from the entire wafer and then prepared for packaging into the final product.

Stages of IC Fabrication

Fabrication of microchips involves five general stages of manufacturing (see Figure 1.6):

- Wafer preparation
- Wafer fabrication
- Wafer test/sort
- Assembly and packaging
- Final test

These five stages are interdependent, with a large infrastructure within the semiconductor company and a support network of industries that provide specialized chemicals and equipment. Companies that operate only in an individual stage (such as a chip company that only fabricates chips) have to meet industrywide standards to ensure that the final microchip meets performance objectives.

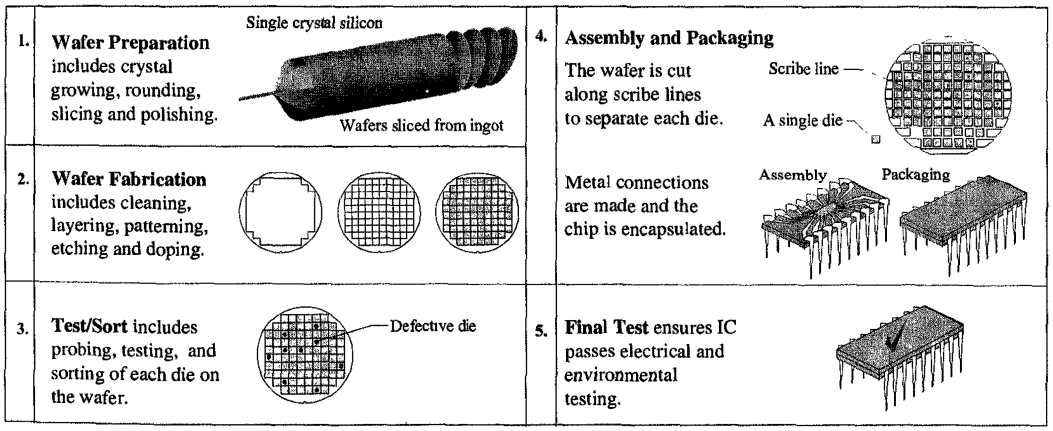


FIGURE 1.6 Stages of IC Fabrication

Wafer Preparation ■ In the first stage, *silicon* is mined from sand and purified. It undergoes specialized processing to create a silicon ingot with the appropriate diameter (see Figure 1.7 on page 8). The silicon ingot is then sliced into thin wafers that are used to fabricate the microchips. Wafers are prepared to special specifications for criteria such as flatness and contamination.

Since the 1980s, most companies that fabricate microchips purchase their wafers from suppliers who specialize in crystal growth and wafer preparation. The industry also produces wafers made from germanium or compound semiconductor materials. These are specialized applications, with the majority of all semiconductor wafers made from silicon.

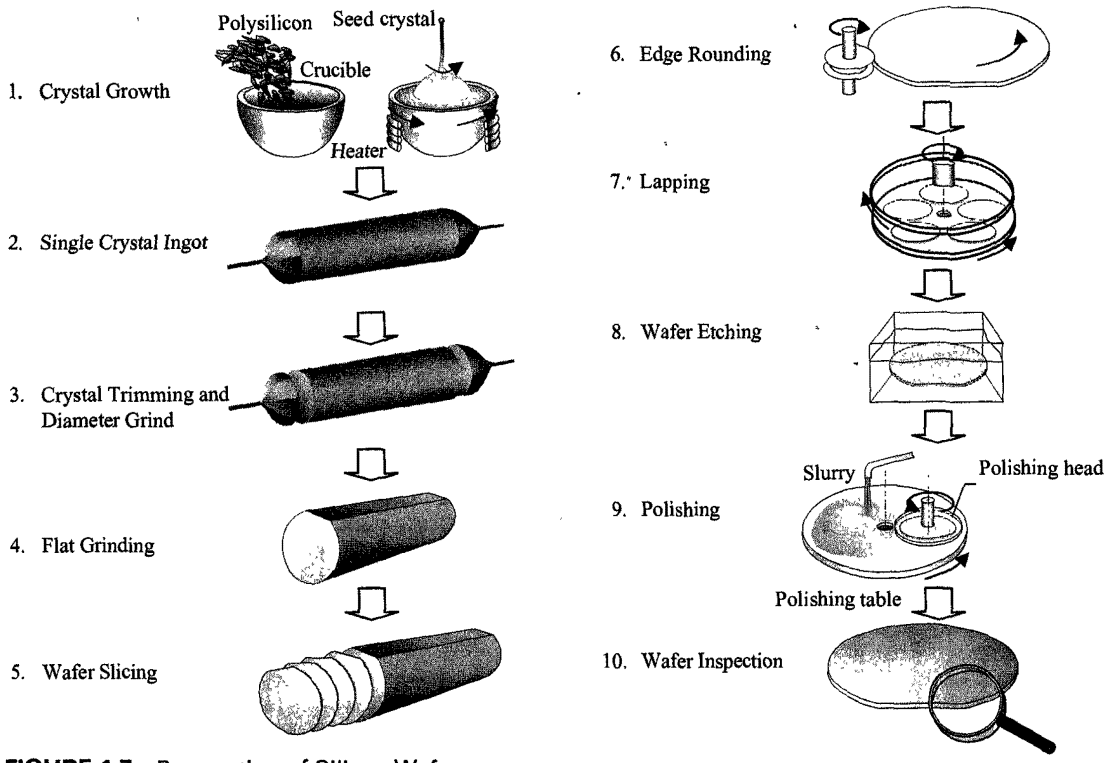
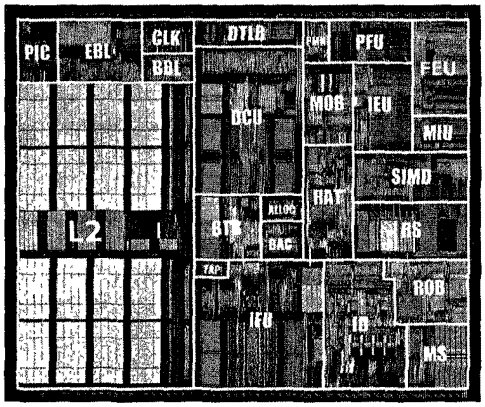


FIGURE 1.7 Preparation of Silicon Wafers
(*Note:* Terms in Figure 1.7 are explained in Chapter 4.)

Wafer Fabrication ■ The fabrication of microchips from wafers is the second stage, referred to as *wafer fabrication.* The bare silicon wafers arrive in the wafer fab, then they go through a variety of cleaning, layering, patterning, etching, and doping steps. The completed wafers have a full complement of integrated circuits permanently etched into each silicon wafer. Other names for wafer fabrication are *microchip fabrication, chip fabrication,* and *chip fab.*

Companies that fabricate chips are both merchant and captive producers. *Merchant chip suppliers* fabricate chips to sell on the open market, such as a chip manufacturer that produces memory chips for customers. *Captive chip producers* fabricate chips to be used in-house in their company's products. An example of a captive producer is a company that makes computers and also fabricates chips for their computers. Some chip makers fabricate chips for in-house use and also sell chips on the open market, while others will fabricate specialty chips and purchase other chips on the open market.

Another type of chip maker is the *fabless company.* This type of company designs chips for a particular market, such as a video microchip, while another chip maker fabricates the chips. Finally,



ULSI Chip
(Photo courtesy of Intel Corporation)

another type of semiconductor manufacturer, the *foundry,* produces chips only for other companies. Semiconductor foundries have become more common since the 1980s, with about 10% of all chips now fabricated at foundries. A major reason for the increase in fabless chip producers and foundries is the high cost of constructing and maintaining a wafer fab. Currently, a wafer fab for high-performance ICs costs around $1.5 to $3.0 billion, with about 75% of the total cost for equipment.

Wafer fabrication of ICs involves the interaction of many complicated process steps using automated equipment to produce the hundreds of millions of devices on one ULSI wafer. Because of the complexity associated with fabricating high-performance ICs, the semiconductor industry has always been at the leading edge of equipment design and manufacturing technology. This innovation has fostered continual improvements in wafer fabrication.

Wafer Test/Sort ■ At the completion of wafer fabrication, the wafer is then sent to the test/sort area where individual die are probed and electrically tested. Defective die are marked for sorting later into acceptable and unacceptable chips. Die that fail the wafer test will not be sent to customers, while die that pass the wafer test will continue through the process.

Assembly and Packaging ■ After wafer test/sort, the wafer enters assembly and packaging to enclose the individual chips in a protective package. The backside of the wafer undergoes a backgrind to reduce the thickness of the substrate. A thick membrane of acetate tape is attached to the backside of each wafer, then the front of the wafer is sliced along the scribe lines with a diamond-tip saw blade to separate the die from each wafer. The sticky acetate holds the silicon chips to keep them from falling off. At the assembly plant, the good die have wire bonds added or have bumps applied to form the attachment to the assembly package. Soon after, the die is hermetically sealed in a plastic or ceramic package. The actual form of the final package varies depending on the type of chip and its application (see Figure 1.8 on page 10). See Chapter 20 for more information on assembly and packaging.

Final Test ■ To ensure a chip's functionality, each packaged integrated circuit is tested to meet the manufacturer's electrical and environmental specifications. After the final test, the chips are delivered to the customer for assembly into the specific application, such as mounting memory components on a circuit board for a personal computer.

SEMICONDUCTOR TRENDS

The rapid technology changes needed to design and fabricate ICs with ULSI integration leads to the continual introduction of new equipment and processes. The semiconductor industry introduces new fabrication technology into the wafer fab every 18 to 24 months. Wafer fabrication technology changes are driven by the customers' needs. Customers require faster, more reliable, and lower cost chips. To achieve this, chip manufacturers have learned to reduce the size of components on a chip. This increases chip speed and reduces power consumption. Chips undergo extensive testing and analysis to verify long-term reliability. The number of chips on a wafer are increased to reduce cost.

There are three major trends associated with improvements in microchip technology:

- Increase in chip performance
- Increase in chip reliability
- Reduction in chip cost

As the different process areas of wafer fabrication are studied in this text, the introduction of new technologies will be discussed based on their contribution toward these trends.

Increase in Chip Performance

The performance of semiconductor microchips has dramatically improved since the SSI era of the early 1960s. A common way to judge chip performance is speed. *Chip speed* will improve if devices are made smaller and packed closer together on the chip because the electrical signal moving through the circuit has less distance to travel. Another way to increase speed is to use materials that

improve the passage of the electrical signal through the device and circuitry on the chip surface. We will study these new materials in later chapters. Microprocessor chip performance is also judged by the number of instructions the chip can perform, as measured in million instructions per second (MIPS). A faster microchip that can perform more instructions per second is beneficial to the customer.

Critical Dimension (CD) ■ The physical dimension of a feature on a chip is referred to as the *feature size*. Another term to describe feature size is *circuit geometry*. Of special note is the minimum feature size on a wafer, known as the *critical dimension,* or *CD*. We will refer throughout the text to the CD as the criteria that defines the level of fabrication complexity (i.e., if you possess the capability to fabricate the CD on a wafer, then you can fabricate all other feature sizes because the dimension is larger and therefore easier to produce). For instance, if the smallest linewidth on a chip is 0.18 μm, then this dimension is the CD (see Figure 1.9). Device CDs have continually shrunk since the beginning of semiconductor fabrication, starting with a CD of about 125 μm in the early 1950s and currently at 0.18 μm and less.[9] The semiconductor industry uses the term *technology nodes* to describe the applicable CD in use during wafer fabrication. The actual and projected industry technology nodes for CDs from 1 μm and below are shown in Table 1.2.

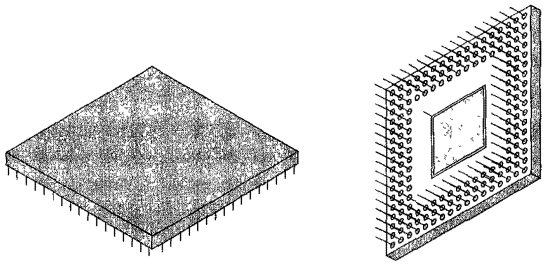


FIGURE 1.8 Sample of Microchip Packaging

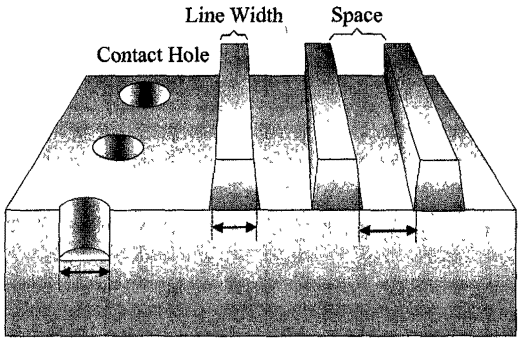


FIGURE 1.9 Critical Dimension

TABLE 1.2 Past & Future Technology Nodes for Feature Critical Dimension (CD)

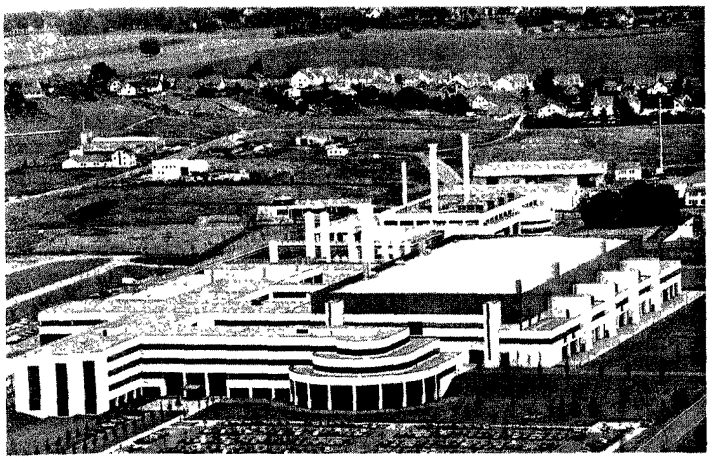
	1988	1992	1995	1997	1999	2001	2002	2005
CD (μm)	1.0	0.5	0.35	0.25	0.18	0.15	0.13	0.10

Adapted from *1999 Roadmap: Solutions and Caveats,* Solid State Technology, May, 2000, p. 77.

The coordinated shrinking of device dimensions on the chip is referred to as *scaling*. It is not acceptable to only reduce one feature on a chip. This would be like designing a large pickup truck with very small tires. Normally, a large vehicle for heavy loads requires large tires. A small car with a light load is acceptable with small tires. The same goes for chip design. All dimensions must be reduced simultaneously, or scaled, in order to optimize electrical performance. For advanced semiconductor manufacturing, device scaling occurs for dimensions in both the vertical and lateral directions.

Components Per Chip ■ Reducing the feature size on a chip permits more components to be fabricated on the silicon wafer. For microprocessors, the number of transistors on the chip surface can illustrate the increased chip integration permitted by the reduction in CD. Since the number of transistors on a chip has dramatically increased over the years of wafer fabrication, there has been an improvement in chip performance (see Figure 1.10).

Moore's Law. In 1964, Gordon Moore, one of the early pioneers in the semiconductor industry and a founder of Intel, predicted that the number of transistors on a chip would double roughly every year.[10] This prediction is known in the industry as *Moore's law* (it was later modified



Wafer Fab
(Photo courtesy of Advanced Micro Devices-Dresden Copyright by S. Doering)

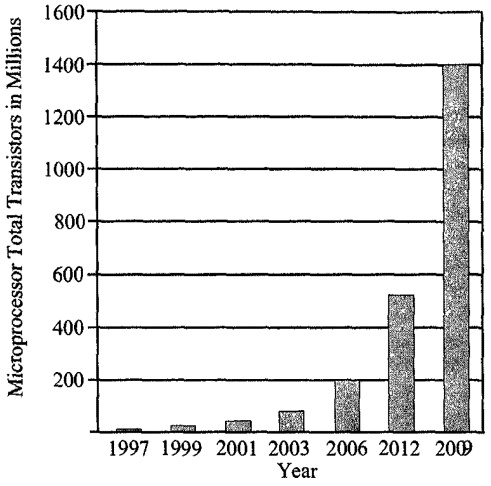


FIGURE 1.10 Increase in Total Transistors/Chip
Redrawn from Semiconductor Industry Association, *The National Technology Roadmap for Semiconductors,* 1997

in 1975 to predict doubling every 18 months). A key contributor to Moore's law is the ability to fabricate wafers with a reduction in the device feature size to achieve a new CD and an increase in the number of transistors on a chip with the introduction of each new product generation. Moore's law for transistors on microprocessors, shown in Figure 1.11 has been surprisingly accurate.

Size reduction in semiconductor devices as predicted by Moore's law is important because it leads to smaller microchip packaging. Smaller microchips fit into small volumes, which leads to smaller commercial products. Semiconductors are the enabling technology for the development of portable electronic products such as the palmtop computer and cellular telephones (see Figure 1.12 on page 12).

Power Consumption ■ Another important aspect of chip performance is the power consumption of the device during operation. Whereas vacuum tubes were power hungry, semiconductor devices consume substantially less power. As device miniaturization proceeds, there is a corresponding reduction in power consumption.[11] Although the number of transistors per chip is rapidly increasing, the chip power consumption grows at a much slower rate (see Figure 1.13 on p. 12). This has become a key performance parameter given the market growth for portable electronic products.

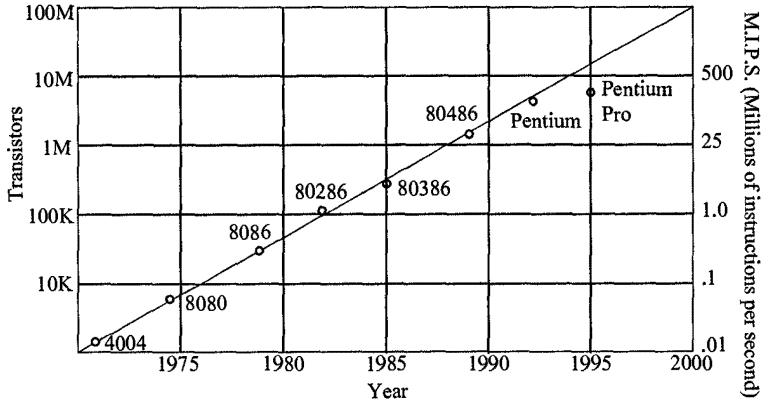


FIGURE 1.11 Moore's Law for Microprocessors
(Used with permission from *Proceedings of the IEEE,* January, 1998, © 1998 IEEE)

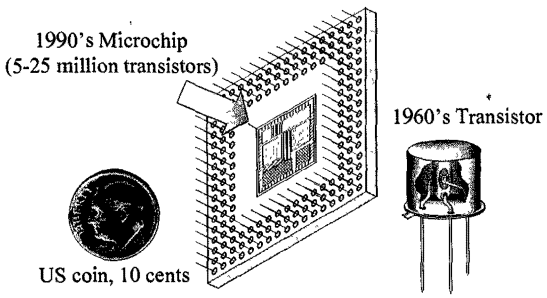


FIGURE 1.12 Size Comparison of Early and Modern Semiconductors

Increase in Chip Reliability

Chip reliability addresses the ability of the chip to function as intended over its expected life. Advances in technology have improved the product reliability of chips (see Figure 1.14). For example, contaminants are controlled through strict application of cleanroom procedures in areas such as particle-free air and control of chemical purity. Fabrication processes are constantly analyzed to improve device reliability. Wafers are monitored and microchips tested to verify acceptable performance. This translates into products that function with low failures during operation.

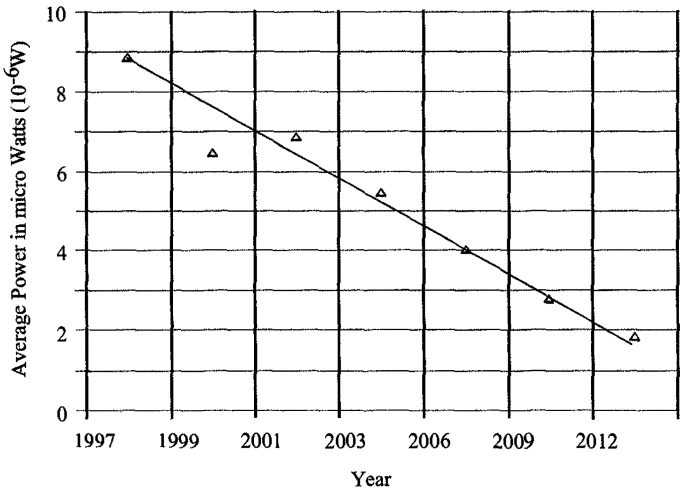


FIGURE 1.13 Reduction in Chip Power Consumption per IC
Redrawn from Semiconductor Industry Association, *The National Technology Roadmap for Semiconductors,* 1997

Reduction in Chip Price

The price of semiconductor microchips has steadily decreased (see Figure 1.15). During the nearly 50 years prior to 1996, the price of semiconductor microchips has decreased by a phenomenal 100 million times.[12] For example, in 1958 a single silicon transistor that was of poor quality cost around $10. Today, $10 might buy a memory chip with over 20 million transistors, an equal number of other components, and the necessary interconnections to make a useful chip.

There are several reasons for the decreasing prices for semiconductor chips. As previously explained, factors such as reducing the feature size and increasing the wafer diameter to put more chips on a wafer are contributors to price reduction. The cost benefits from this decrease in CD are substantial. For example, if a semiconductor manufacturer decreased the CD on an 8-inch wafer from 0.35 μm to 0.25 μm in 1997, they were able to increase the number of chips per wafer from 150 to 275. In other words, they produced almost twice as many chips per wafer at nearly the same manufacturing cost. This cost benefit was achieved because the reduced CD permitted millions of circuit lines to be compressed into a smaller area.

Another reason for price reductions has been the large market growth for semiconductor products. This growth requires increased volume from chip fabrication companies, which introduces economies of scale for manufacturing. The introduction of manufacturing improvements to

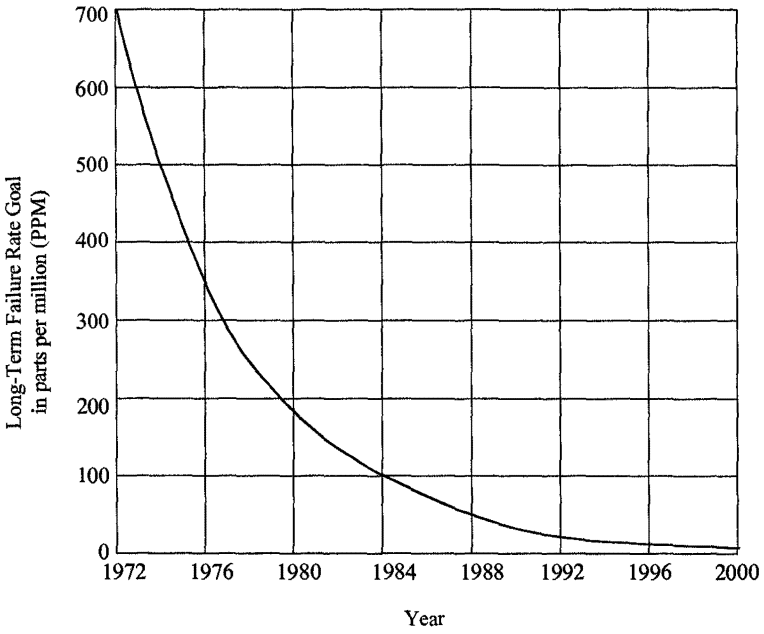


FIGURE 1.14
Reliability Improvement of Chips

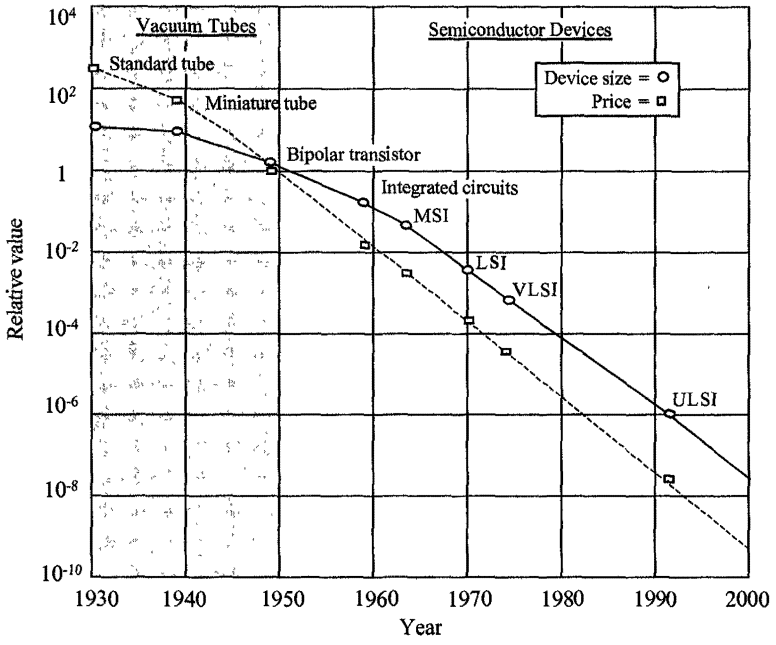


FIGURE 1.15 Price Decrease of Semiconductor Chips Redrawn from C. Chang and S. Sze, *ULSI Technology,* (New York: McGraw-Hill, 1996), p. xxiii.

the equipment and process used to fabricate microchips has also reduced costs. Work teams have an important role in improving manufacturing technology and cost. They work firsthand with the chip fabrication equipment and contribute to improving manufacturing productivity. The efforts of manufacturing teams make a direct difference in lowering the cost to manufacture chips.

THE ELECTRONIC ERA

The semiconductor industry has undergone intensive growth since the invention of the first transistor at Bell Labs in 1947. It is beneficial to review the major industry changes since the beginning of solid state electronics to gain insight from the foundations of modern wafer fabrication.

The 1950s: Transistor Technology

The 1950s were a rich era in the development of different types of semiconductor transistor technology. Many different types of transistors were developed at Bell Labs and the early semiconductor efforts in northern California and Texas. The first transistors were developed from germanium semiconductor material. In 1954 the first manufacturable silicon transistor was made by Gordon Teal at Texas Instruments.[13]

With the invention of the IC in 1959, the industry was poised to solidify transistor technology as the replacement for vacuum tubes and to develop new market applications. The basic elements needed to commercialize the transistor were in place. The semiconductor industry was growing in high technology centers such as Silicon Valley in northern California; Route 128 near Boston, Massachusetts; and Texas Instruments in Dallas, Texas. This concentration of ideas and technology was important for the development of semiconductor technology and the new manufacturing techniques necessary to move wafer fabrication out of the laboratory processes of the 1950s into production processes of the 1960s. The growing dominance of semiconductors and silicon in all aspects of society gave credence to the expression, "the silicon age."

The 1960s: Process Technology

In the 1960s, the semiconductor industry entered an era oriented toward solving the basic problems of producing semiconductor ICs. This was the beginning of integration and the SSI era. Semiconductor manufacturers were proliferating. The market for semiconductor ICs was growing rapidly, with sales exceeding $1 billion in 1961. The subsequent growth of the many chip manufacturers served to increase IC performance and reduce cost. Because of process commonality across semiconductor companies, special supplier industries developed to provide the chemicals and equipment needed for wafer fabrication.

There was a proliferation of engineers wanting to exploit the new semiconductor technology, which led to the formation of many new high-tech companies in the 1960s. Signetics was founded in 1961 by a group of Fairchild engineers to commercialize ICs. Robert Noyce, Gordon Moore, and Andrew Grove left Fairchild to found Intel in 1968. Jerry Sanders and other scientists, also from Fairchild, founded Advanced Micro Devices in 1969. There were also some large captive supplier companies in the IC business, such as IBM and Digital Equipment Corporation, and midsize companies such as Hewlett-Packard. The exploding demand for semiconductor products helped the many small start-up companies overcome the hurdles of constant price reductions, rapid company growth, and financial challenges of starting new companies.

The 1970s: Competition

The early 1970s was the period of medium-scale integration (MSI) for chip design. The manufacturing processes were largely manual operations based on batch processing. A typical wafer fab was fortunate to start a new product with a yield of 5 or 10% (yield is the percent of acceptable die produced at the completion of the process) and after an intense effort possibly raise it to 30%.[14]

Microprocessors were invented separately in the early 1970s at both Texas Instruments and Intel, and with its wide acceptance in market applications came the need for more chip integration. LSI-level integration remained only a few years and then was quickly replaced by VLSI, the integration standard by the end of the 1970s.

The rapidly changing semiconductor industry became chaotic. Much of the equipment and many of the processes used in the wafer fabs were developed by the same companies that manufactured the semiconductor devices, with few industrywide standards. This created inefficiencies for manufacturers and suppliers. Asia emerged as a formidable competitor, with Japan as a powerhouse of semiconductor innovation and manufacturing. Asian electronic giants challenged

American dominance in semiconductors. By 1979, Japan had captured more than 40% of the world's demand for memory microchips.[15]

As the demand for more complex chips grew, equipment technology changed from the manual tools of the 1960s to semiautomatic operation with buttons for operator control and onboard solid-state controllers. Equipment suppliers also encountered stiff competition from Japan and Asia for key semiconductor tools.

The cost of constructing a wafer fab became extremely expensive. Control of wafer contamination levels became critical for shrinking device feature sizes, requiring special wafer fab cleanliness standards that exceeded nearly every other industry. Special purity requirements were needed for water, air, and the many chemicals and gases used during fabrication. By the end of the 1970s, the cost of a wafer fab was around $30 million and rising. Market demand and the need to continually replace equipment to maintain the advancing technology encouraged the construction of new fabs.

The industry attempted to form collaborative ventures to address the disorder. In 1970, SEMI (Semiconductor Equipment and Materials Institute) was founded to standardize and promote equipment, materials, and services in the industry. The Semiconductor Industry Association (SIA) was started in 1977, with leadership by Robert Noyce. The goal was to have more industry coordination about mutual problems created by rapid growth.

The 1980s: Automation

By the 1980s, the widespread acceptance of solid-state electronics in society was established, driving the growth of chip fabrication in the semiconductor industry. This was the VLSI era, with a high degree of component integration on a chip. The growth of the personal computer industry fueled hardware and software demand. At the same time, the industry faced competitive pressures by Japanese IC manufacturers that expanded their manufacturing capacity while at the same time driving yield and quality to unprecedented levels. U.S. semiconductor companies were clearly shaken by these developments and reached near panic because of their inability to compete. By the mid-1980s, Japan almost totally dominated the fast growing and technologically demanding DRAM (dynamic random access memory) market segment.[16]

The semiconductor industry formed SEMATECH in 1987, under the guidance of the Department of Defense, with Robert Noyce as its first Chief Executive Officer. The goal was to develop industrywide policies with respect to the specification and evaluation of manufacturing equipment. These organizations were formed, in part, because Japanese competition threatened the very existence of U.S. semiconductor manufacturers.[17] Almost simultaneously SEMI/SEMATECH was formed of U.S. equipment and materials suppliers in support of SEMATECH's mission. The names of the consortiums would be changed in 1999 to International SEMATECH and the Semiconductor Industry Suppliers Association, respectively.

Improvements in semiconductor equipment, manufacturing efficiency, and product quality were emphasized in U.S. companies. Fabrication tools were automated to include all significant wafer processing steps. The intent was to largely remove the operator from the process, especially since the human was a major source of contamination in the cleanroom.

The industry was able to maintain the predicted growth of Moore's law by means of ongoing chip design changes and reduction in chip feature sizes. These design changes presented manufacturing challenges that led to the development of sophisticated processes. The equipment and cleanroom controls necessary to fabricate the new ICs escalated the cost of building wafer fabs, which rose to nearly $1 billion by the late 1980s (see Figure 1.16 on page 16). To overcome this high cost of entry into new technology with new fabs, many U.S. semiconductor companies entered into joint venture agreements with Japanese and European chip manufacturers to share the cost of constructing wafer fabs.

The 1990s: Volume Production

Feature size dimensions shrunk below 1 μm for production chips during the late 1980s and early 1990s. By the end of the 1990s, the minimum feature size was at 0.18 μm. Microchip design experts predict that at some point there will be a physical limitation to further reductions in feature size. At this time, there is no barrier anticipated until the CD goes well below 0.1 μm. CDs

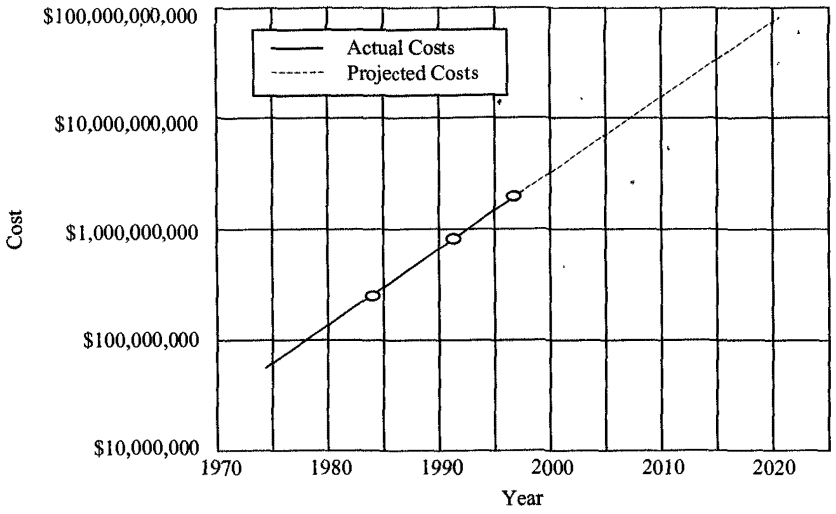


FIGURE 1.16 Start-Up Cost of Wafer Fabs (Used with permission from *Proceedings of IEEE*, January, 1998, © 1998 IEEE)

continue to shrink because of the cost pressures for manufacturing efficiency resulting from putting more chips on the same wafer.

Submicron geometry processing ushered in the ULSI era of integration, with high-performance ICs containing 10 million transistors or more. The two types of chips with the most advanced technology are microprocessors and memory chips. Highly integrated chips require multiple layers of circuitry to interconnect the devices (up to eight layers and increasing) and up to 450 or more process steps to fabricate the chips.

The semiconductor industry has become increasingly competitive in the 1990s. To survive in the worldwide chip market, it is critical for manufacturers to produce complex, high quality chips in record time. Delivery of the right microchip to the market has a small window of opportunity based on customers and their need for advanced technology. Semiconductor manufacturers try to be the first (or nearly the first) with the new technology. If a company misses this product window, then it risks spending substantial money to develop the chip technology without realizing any sales potential.

Semiconductor equipment is highly automated. Advanced material handling systems move wafers between workstations with no required human intervention. Expert software systems control nearly all equipment functions, including troubleshooting diagnostics. Technicians and engineers intervene to download product recipes into the equipment software database and interpret software diagnostic commands to take corrective action for equipment repair.

CAREERS IN SEMICONDUCTOR MANUFACTURING

The career paths in semiconductor manufacturing are grouped into three main areas: technician, engineering, and management. The choice of a particular career path generally depends on an individual's technical knowledge, education, and personal goals. Examples of different career paths are shown in Figure 1.17.

Technician

The technician has the technical skills to operate, troubleshoot and maintain the advanced equipment used to fabricate microchips in a production environment. Equipment operation requires knowledge to set up and operate the automated equipment so that it functions properly in the manufacturing process. Most equipment operation is done through software interfaces to the computer-controlled equipment. Troubleshooting equipment problems requires expertise in equipment systems (e.g., electronics, mechanical systems, motors, etc.) with the ability to interpret software diagnostic routines. Equipment maintenance involves extensive training and knowledge of electrical, mechanical, chemical, and computer equipment systems. An equipment technician often performs the troubleshooting and maintenance of wafer fab equipment.

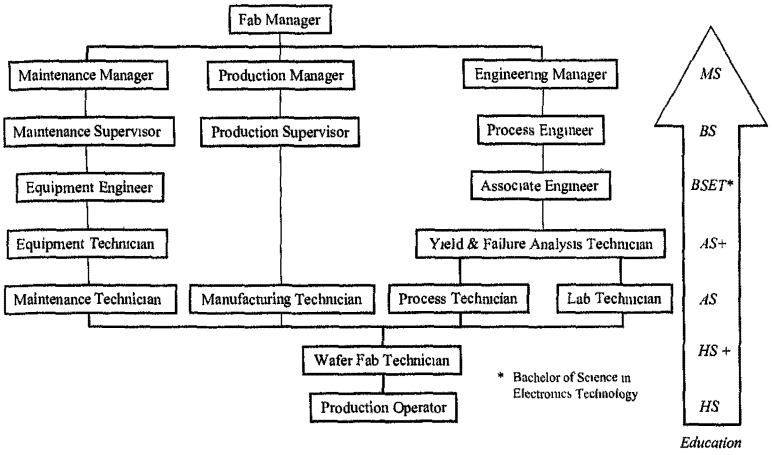
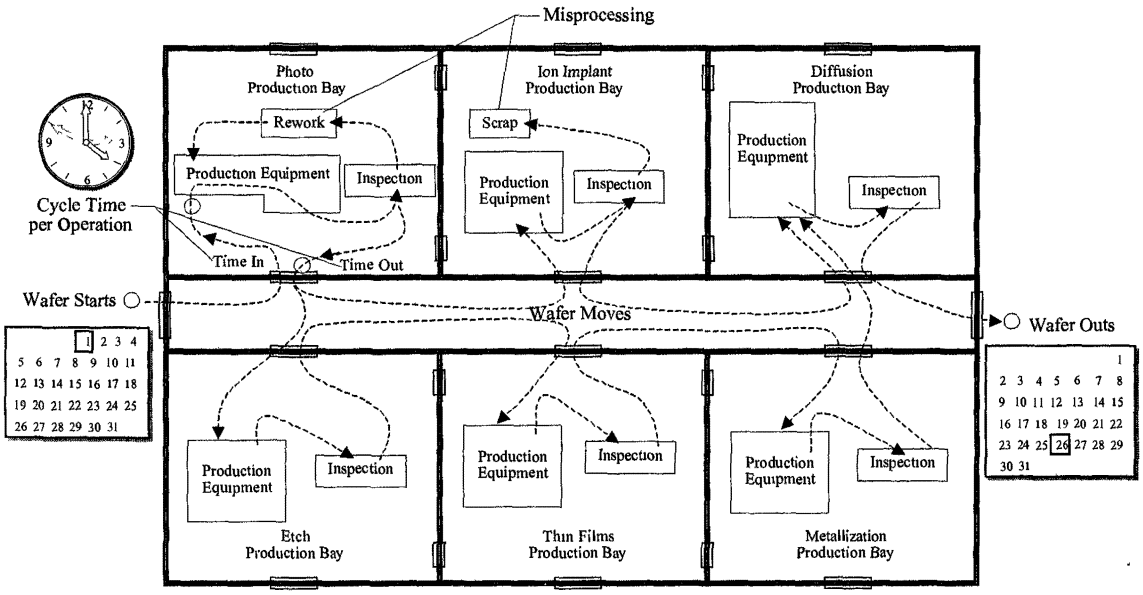


FIGURE 1.17 Career Paths in the Semiconductor Industry

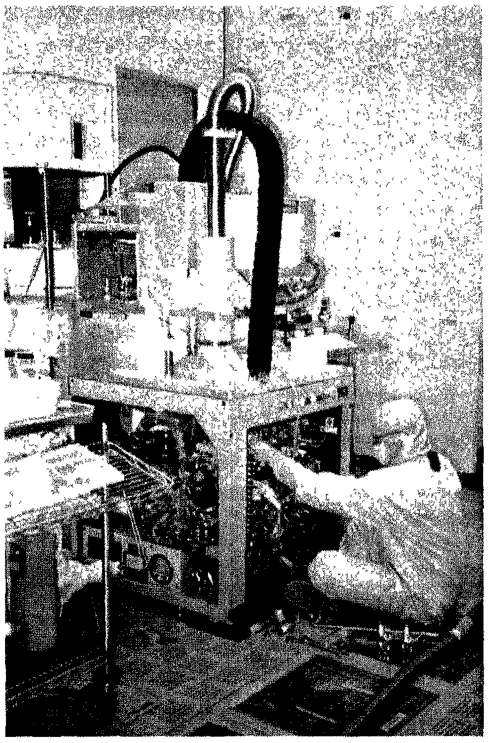
Technicians must be capable of performing their technical role as a manufacturing team member, while contributing to the productivity objectives of the wafer fab (see Figure 1.18). Technicians are knowledgeable about safety procedures for chemicals and equipment. As equipment automation continues toward more integrated equipment across different operations, the skills of the technician in the wafer fabs must also advance.

Technicians in a wafer fab are referred to by different names at the various chip manufacturers, such as a wafer fab technician (WFT), self sustaining technician (SST), manufacturing association (MA), or manufacturing technician (MT). Technicians are also used in the semiconductor industry, such as for field service representatives, to install and service sophisticated wafer fab equipment. The skills required of technicians include computer usage, knowledge of equipment procedures and support, analytical skills to solve problems for process improvement, and critical thinking skills for decision-making. Due to the advanced sophistication of wafer fab equipment, it is clear that people who merely operate equipment are in less demand in the wafer fab. Furthermore, humans create too many problems with respect to wafer contamination, manufacturing repeatability, and inefficiency.

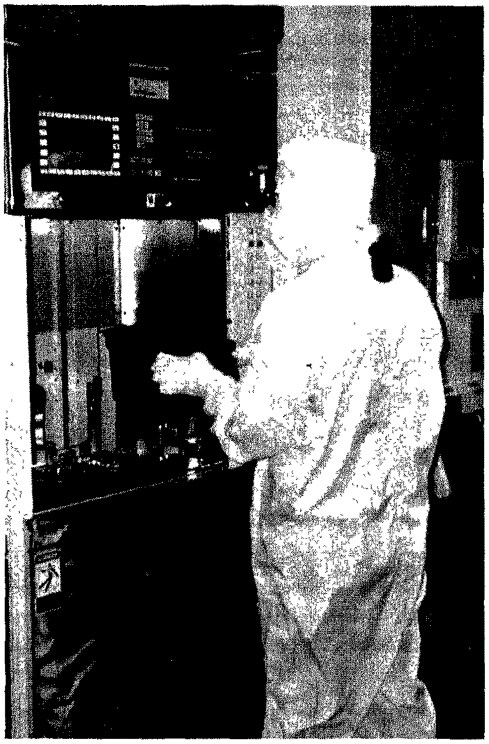


Production Cycle Time = (Date and Time of Wafer Start) - (Date and Time of Wafer Out)
Wafer Outs = Wafer Starts - Wafers Scrapped
Operator Efficiency = Theoretical Cycle Time / Actual Cycle Time

FIGURE 1.18 Productivity Measurements in a Wafer Fab



A. Equipment Technician in a Wafer Fab



B. Technician in a Wafer Fab
(Photos courtesy of Advanced Micro Devices)

Job Descriptions

Within each of the three major career paths, there are various options available for different jobs. A description of some different technician, engineer and management positions in the wafer fab is given.

- Wafer Fab Technician: The wafer fab technician is responsible for operating the wafer fab equipment. The wafer fab technician performs some equipment maintenance and basic troubleshooting of process and equipment. Semiconductor companies often prefer high school or some college (e.g., a certificate). In some cases, a two-year associate of applied science degree may be desirable.
- Equipment Technician: The equipment technician troubleshoots and maintains advanced equipment systems so that the equipment performs correctly during wafer manufacturing. This position typically requires at least a two-year associate of applied science degree and extensive hands-on training.
- Equipment Engineer: The equipment engineer specifies the equipment design parameters and optimizes equipment performance for wafer production. This career typically requires at least a four-year engineering degree.
- Process Technician: The process technician supports the production equipment and process engineering teams by troubleshooting process-related problems. A two-year associate of applied science degree or four-year engineering technology degree is helpful; however, good writing, math, science, and analytical thinking skills are necessary.
- Process Engineer: The process engineer analyzes the performance of the fabrication process and equipment to determine optimum parameter settings. This position typically requires a four-year engineering or science degree.
- Field Service Representative: A field service representative in a wafer fab installs the fabrication equipment. The field service representative also performs equipment maintenance, diagnostics, and repairs to ensure the equipment is production-worthy. This position requires at least a two-year associate degree or four-year engineering technology degree.
- Lab Technician: Lab technicians can work in development laboratories setting up and running experiments. This job skill usually requires some experience and people who are self-motivated.

- Yield/Failure Analysis Technician: These technicians perform the tasks associated with defect analysis, such as preparing the materials to be analyzed and operating analytical equipment to determine the source of problems that arise during wafer fabrication.
- Yield Improvement Engineer: The yield improvement engineer collects and analyzes yield and test data to improve wafer fabrication performance. This position often requires a four-year science or engineering degree.
- Facilities Technician: Facilities technicians support the facility equipment and utilities required in a wafer fab, including chemical management and cleanroom utilities.
- Facilities Engineer: This engineer provides engineering design support for the wafer fab utility infrastructure for chemicals, clean air and equipment that are used in the facility. The position typically requires a four-year engineering degree.
- Supervisor/Manager: Supervisory and management positions in a wafer fab combine technical skills with people management skills to achieve the organizational goals of the company. These positions often will require either a four-year science, engineering, or management degree or an additional two years of college after the associate degree.

SUMMARY

The semiconductor industry is a major part of the high tech industry in the U.S. The invention of the solid-state transistor launched the semiconductor industry in 1947, followed by the introduction of silicon and the development of the integrated circuit (IC). The IC combines multiple components on one chip to increase chip performance and lower cost. There are IC integration eras, with the most recent era being ULSI (ultra large scale integration). IC fabrication occurs on a wafer, with the devices on the top of the silicon and layers of circuitry deposited on the substrate. There are five stages of microchip fabrication: wafer preparation, wafer fabrication, wafer test/sort, assembly and packaging and final test. The three major trends for improvement in microchip technology are: increase in performance, increase in reliability, and reduction in cost. Chip speed is important for performance and is improved by reducing the critical dimension (CD), or minimum feature size, on a chip. Moore's law predicts that the number of components on a chip will double every 18-24 months, which has remained relatively true since the 1960s. The reliability of chips has improved while the price of chips has substantially decreased since 1947. The semiconductor industry has become increasingly competitive and the wafer fab technician plays an important role in a manufacturer's ability to fabricate microchips.

KEY TERMS

semiconductor
wafers
wafer fab
microchips (chips)
solid-state transistor
planar technology
integrated circuit (IC)
devices
wafer
die
substrate
wafer fab
fab
silicon
wafer fabrication
microchip fabrication
chip fabrication
chip fab
merchant chip suppliers
captive chip producers
fabless company
wafer sort
packaging
chip speed
feature size
circuit geometry
critical dimension (CD)
technology node
scaling
Moore's law
wafer fab technician (WFT)

REVIEW QUESTIONS

1. How big is the high-tech industry relative to the U.S. economy? What has been a major factor in helping the semiconductor industry become so large?
2. List four different industries from the first half of the twentieth century that contributed to the development of the semiconductor industry.
3. When, where, and by whom was the solid-state transistor invented?
4. What is an integrated circuit (IC)? When was it invented, and by whom?
5. List the five integration eras, provide the time period for each era, and give the number of components on a chip in each era.
6. What is a wafer? What is a substrate? What is a die?
7. Give a short description of a wafer fab.
8. List the five major stages of IC fabrication, and give a short description of each stage.
9. Describe merchant chip supplier, captive chip producer, fabless company, and foundry.
10. List the three major trends associated with improvement in microchip fabrication technology, and give a short description of each trend.
11. What is the chip critical dimension (CD)? Why is this dimension important?
12. Describe scaling and its importance in chip design.
13. What is Moore's law and what does it predict? Has this law been basically true?
14. By what factor have chip prices reduced since 1947? Give two reasons for this change.
15. Provide a short description of how the electronic era developed from 1950 through 2000.
16. Describe the responsibilities of a wafer fab technician and an equipment technician.
17. List and give a short description of eight different job descriptions in the semiconductor industry.

SELECTED INDUSTRY WEB SITES

Advanced Micro Dvices	http://www.amd.com
AT&T Tech. History	http://akpublic.research.att.com/
Semiconductor International Magazine	htt;p://www.semiconductor.net/
Fairchild Semiconductor	http://www.fairchildsemi.com/
IBM Microelectronics	http://www.chips.ibm.com/
Intel	http://www.intel.com/
International SEMATECH	http://www.sematech.org/public/index.htm
MATGEC, Maricopa Advanced Technology Education Center	http://matec.org/
Motorola Semiconductor	http://mot-sps.com/
Mitsubishi	http://www.mmc-sil.com/
National Semiconductor	http://www.national.com/
NIST, National Institute of Standards and Technology	http://www.nist.gov/
Selete, Semiconductor Leading Edge Technologies, Inc.	http://ww.selete.co.jp
SEMI, Semiconductor Equipment and Materials International	http://www.semi.org/
Semiconductor Search Engine	http://www.semiseek.com/
SIA, Semiconductor Industry Association	http://www.semichips.org/
SISA, Semiconductor Industry Suppliers Association	http://www.sisa.org/
Solid State Technology Magazine	http://sst.pennwellnet.com/home/home.cfm
Texas Instruments	http://www.ti.com/

REFERENCES

1. *Chip Scale Review,* Volume 3, Number 4 (July-August 1999): p. 25.
2. Historical information generally taken from F. Seitz and N. Einspruch, *Electronic Genie: The Tangled History of Silicon* (Urbana: University of Illinois Press, 1998), p. 52.
3. F. Seitz and N. Einspruch, *Electronic Genie,* p. 52.
4. P. Bondyopadhyay, "In the Beginning," *Proceedings of the IEEE* 86, no. 1 (January 1998): p. 63.
5. J. Pierce, "The Naming of the Transistor," *Proceedings of the IEEE* 86, no. 1 (January 1998): p. 37.
6. Texas Instruments, *The Chip That Jack Built Changed the World,* Texas Instruments, (September 1997).
7. G. Moore, "The Role of Fairchild in Silicon Technology in the Early Days of 'Silicon Valley,'" *Proceedings of the IEEE* 86, no. 1 (January 1998): p. 59.
8. Ibid., p. 54.
9. Ibid., p. 56.
10. G. Moore, "Cramming More Components onto Integrated Circuits," *Proceedings of the IEEE* 86, no. 1 (January 1998): p. 84. Reprinted from *Electronics,* April 19, 1965.
11. C. Chang and S. Sze, *ULSI Technology* (New York: McGraw-Hill, 1996), p. xxi.
12. Ibid., p. xix
13. I. Ross, "The Invention of the Transistor," *Proceedings of the IEEE* 86, no. 1 (January 1998): p. 17.
14. F. Seitz and N. Einspruch, *Electronic Genie,* p. 94.
15. C. Melliar-Smith, M. Borrus, D. Haggan, T. Lowrey, A. Vincentelli, W. Troutman, "The Transistor: An Invention Becomes a Big Business," *Proceedings of the IEEE* 86, no. 1 (January 1998): p. 93.
16. Ibid., p. 94.
17. F. Seitz and N. Einspruch, *Electronic Genie,* p. 238.

CHAPTER 2

CHARACTERISTICS OF SEMICONDUCTOR MATERIALS

Understanding the structure and bonding of atoms provides the key to knowing how silicon performs its vital role as a semiconductor. Once unlocked, this knowledge provides the foundation for understanding how the simplicity of a semiconductor device can be a part of the complex world of the microchip.

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Describe the atom, including the valence shell, band theory, and ions.
2. Interpret the periodic table of the elements with regards to main-group elements, and explain how ionic and covalent bonds are formed.
3. State the three classes of materials, and describe each one with regards to current flow.
4. Explain resistivity, resistance, and capacitance, and discuss their importance to wafer fabrication.
5. Describe pure silicon and give four reasons why it is the most common semiconductor material.
6. Explain doping and discuss how the trivalent and pentavalent dopant elements make silicon a useful semiconductor material.
7. Discuss the difference between p-type (acceptor) silicon and n-type (donor) silicon, describe how silicon resistivity changes with the addition of a dopant, and explain the pn junction.
8. Discuss alternative semiconductor materials, with emphasis on gallium arsenide.

INTRODUCTION

The most important semiconductor material in use today is silicon. Devices are fabricated in silicon and function as a semiconductor because of the unique properties of this material. In this chapter, we will first review some physical and chemical concepts associated with the individual atom. We then will use this information to study properties of silicon material to learn its unique physical and chemical structure and how it permits a semiconductor to function. This chapter will establish basic material and semiconductor concepts that will be used throughout the different process steps in wafer fabrication.

ATOMIC STRUCTURE

Matter is everything that has shape, size, and occupies space in the universe. If we take a material, such as a grain of sand, and continually divide it into smaller and smaller particles, we eventually arrive at the smallest particle, an atom. In the atomic model, atoms consist of three types of particles: the neutral neutrons, the positively charged protons that together make up the nucleus of the atom, and the negatively charged electrons that orbit the nucleus (see Figure 2.1). The number of protons in the atom equals the number of electrons, which makes the atom electrically neutral.

An *element* is a substance formed from one type of atom and is the simplest substance with unique physical and chemical properties. A *molecule* consists of two or more atoms in a structure that are chemically bound together, behaving as an independent unit. A *compound* is a substance composed of two or more atoms (similar or dissimilar) that are chemically bound together and have formed a new substance which has different properties from its individual component atoms.

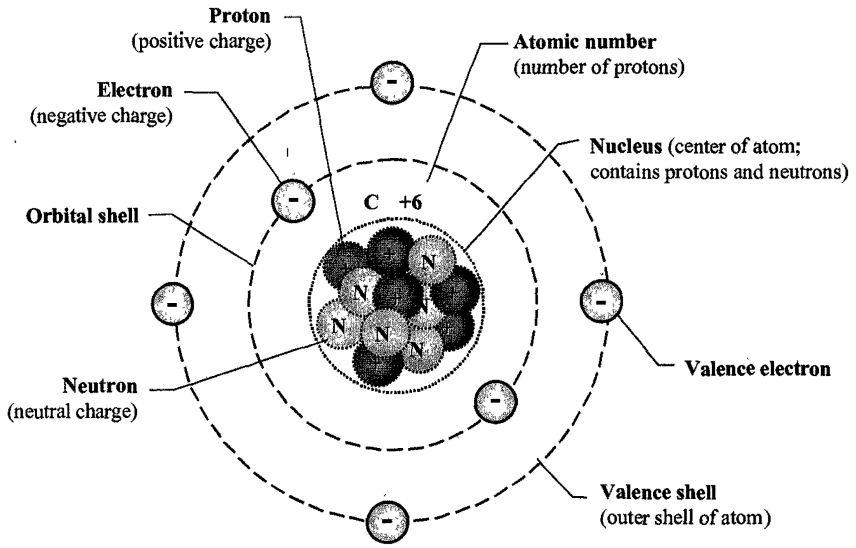


FIGURE 2.1 Elementary Model of the Carbon Atom

Electrons

The basic law of attraction and repulsion that occurs at the microscopic level states that like charges repel and unlike charges attract. Orbiting negative electrons are distributed in a space cloud around the nucleus and held in place by their attraction toward the positive nucleus. Individual electrons occupy orbitals within seven distinct shells. In a hydrogen atom, for example, each shell corresponds to a specific energy level and is identified using the letters K through Q (see Figure 2.2). For all other atoms, referred to as many-electron atoms, the orbital energies within each shell are not the same.

Electron Energy ■ The unit of energy at the atomic level is the electron volt (eV). It represents the kinetic energy (energy associated with motion) gained by an electron in passing from a point of low potential to a point 1 volt higher in potential. One electron volt is equal to 1.6×10^{-19} joules of energy. The electron volt is used to describe electron energy in different processes of semiconductor manufacturing (e.g., ion implant energy levels, as described in Chapter 17).

Valence Shells ■ The outermost electron shell for a given atom is the *valence shell.* Valence electrons are found in the valence shell and have considerable influence on the chemical and physical properties of the atom and exist at the highest energy level for a given atom. The number of

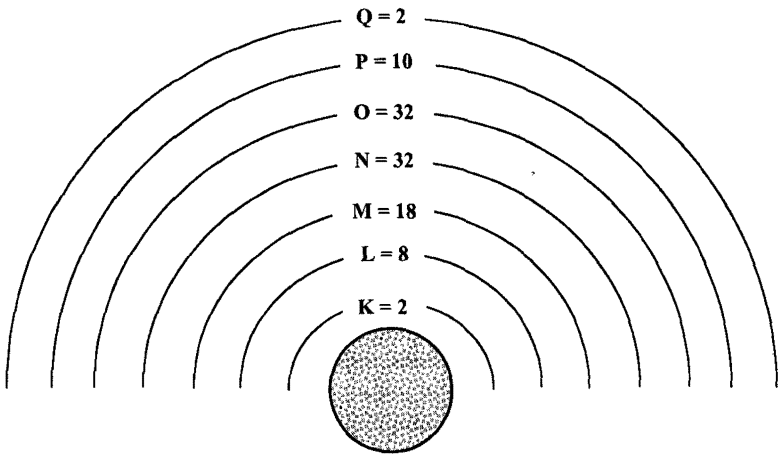


FIGURE 2.2 Electron Shells in Hydrogen Atom

electrons in a valence shell varies depending on the shell level, K to Q. Atoms with one electron in a valence shell that permits eight electrons easily give up this electron. On the other hand, atoms with seven electrons in a valence shell have an affinity for electrons and readily accept one electron to fill the valence shell. The electrons and valence shells for sodium and chlorine are shown in Figure 2.3.

Energy-Band Theory of Solids ■ Energy-band theory explains how electrons change orbital levels in a solid material. The valence electrons are in the valence band. There is a band gap between the valence band and the conduction band (see Figure 2.4 on page 24). The band gap has very high energy levels in some materials that creates a forbidden gap (usually > 2 eV). These materials are referred to as *insulators* because it is hard to move an electron from the valence band to the conduction band. Other materials, called *conductors,* have a valence band that overlaps the conduction band, requiring very little energy for electrons to move into the conduction band. A third class of materials has a band-gap energy level found somewhere between insulators and conductors. These materials are referred to as *semiconductors.* The forbidden gap of semiconductors is moderate. The band-gap energy of silicon is 1.11 eV.

Ions ■ Ions are formed when an atom gains or loses one or more electrons. An atom becomes positively charged if it loses an electron and negatively charged if it gains an electron. Ions with opposite charges are mutually attracted and can form a chemical bond as an ionic compound.

An example of an ionic compound is common table salt, sodium chloride (see Figure 2.5 on page 24). A sodium atom, which is a metal, is initially neutral because it has the same number of

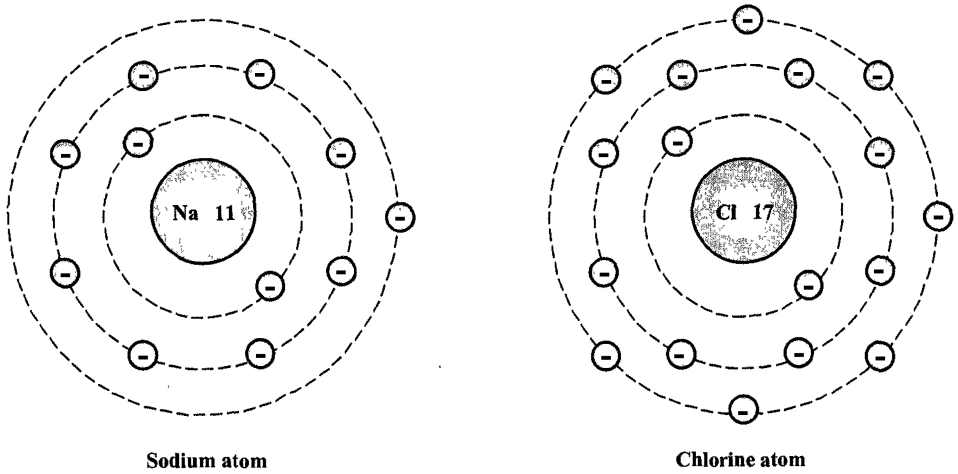


FIGURE 2.3 Electron Shells for Sodium and Chlorine Atoms

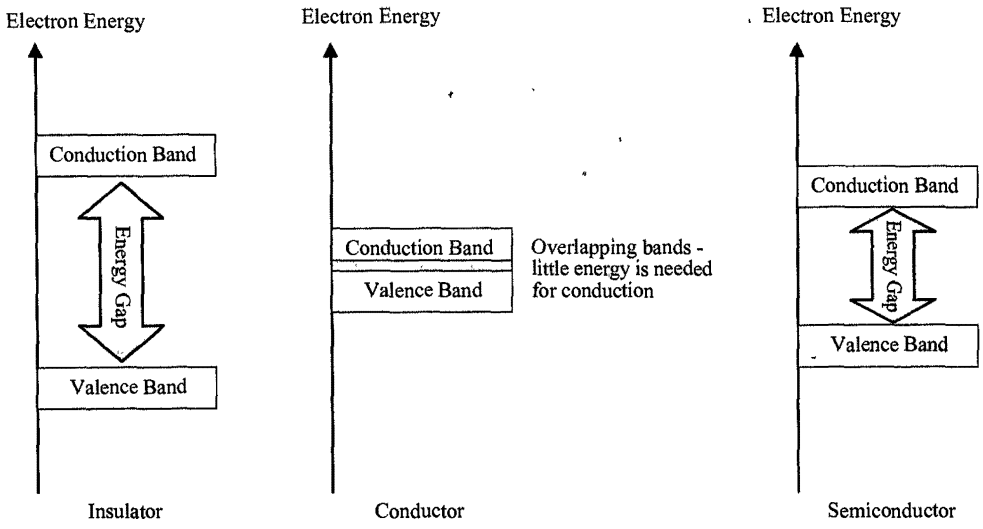


FIGURE 2.4 Energy Band Gaps

protons as electrons. If it loses one electron, it becomes positively charged as Na^+ and is referred to as a *cation.*[1] A chlorine atom, which is a nonmetal, gains an electron and becomes negatively charged as Cl^-, referred to as the *anion.* The Na^+ and Cl^- are attracted to one another due to opposite charges and form an ionic compound.

Many gases exist as molecules. *Ionization* is the removal of electrons from atoms, creating positively charged atoms or molecules. Ionization is used in many process areas in semiconductor fabrication. Once gas particles are charged by ionization, the gas flow and motion of atoms can be controlled by the use of electrostatic and magnetic fields.

THE PERIODIC TABLE

The periodic table lists all elements known at this time (see Figure 2.6). It is organized in such a manner because of certain periodic patterns of behavior among the elements.

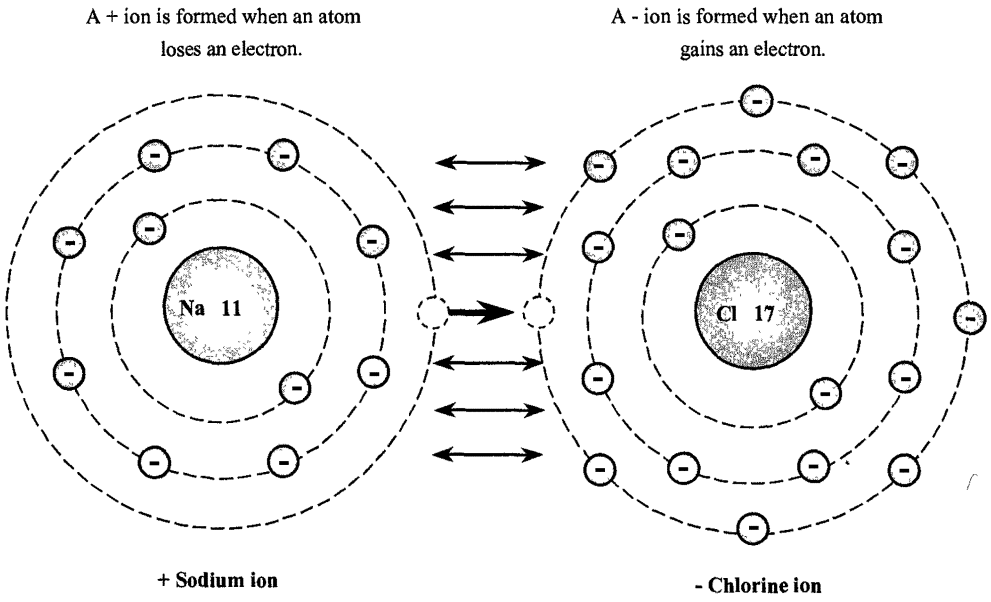


FIGURE 2.5 Sodium Chloride

FIGURE 2.6 The Periodic Table of the Elements

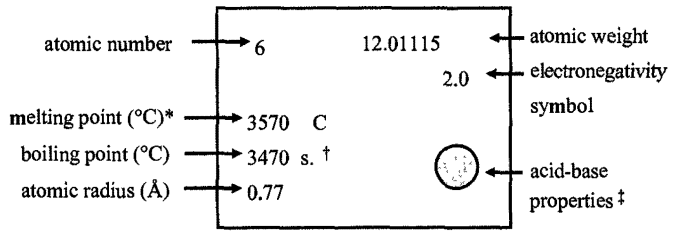
The periodic table has a box for each element (see Figure 2.7). Each element box on the periodic table provides information about the element, including the following:

Atomic symbol: The symbol for each element (e.g., *Si* is silicon).

Atomic number: Equals the number of protons in the nucleus. All atoms of a particular element have the same atomic number (and therefore the same number of electrons).

Atomic mass number: The sum of the protons and neutrons in an atom. Isotopes of the same element have the same number of protons but different mass numbers (therefore, each one has a different number of neutrons).

Atomic weight: Also referred to as atomic mass. This figure is the average of the masses of the naturally occurring isotopes of an element and is weighted according to the abundance of these isotopes. One *atomic mass unit* (*amu*) is a relative measure of mass and is exactly equal to $^1/_{12}$ the mass of a carbon-12 atom.



* Based on carbon-12. ( ) indicates the most stable or best-known isotope.

† s. indicates sublimation

‡ For representative oxides of group. Oxide is acidic if color is red, basic if color is blue, and amphoteric if both colors are shown. Intensity of color indicates relative strength.

FIGURE 2.7 Element Box of the Periodic Table

We will be concerned primarily with the main-group elements of the periodic table found in the columns with group numbers from IA through VIIIA. In the periodic table, the group number is given for each column and represents the number of valence-shell electrons. For instance, Boron (B) is in Group IIIA (found in column IIIA) and therefore has three valence electrons. The group number characteristics of elements commonly used in the semiconductor industry are outlined in Table 2.1.

Ionic Bonds

An ionic bond is formed when valence-shell electrons are transferred from the atoms of one element to another. Unstable atoms easily form ionic bonds. An example is sodium chloride. Sodium (Na) is located in Group IA, indicating there is one electron in its valence shell. Na is an unstable element that is highly corrosive. It is carefully controlled in wafer fabrication to avoid device contamination (see Chapter 6). Chlorine (Cl) is found in Group VIIA and has seven valence-shell electrons. This atom lacks one electron for a fully occupied valence shell; therefore, it also is unstable. Because of their instability, these two elements (Na and Cl) have an affinity for each other. Na readily gives up its valence-shell electron to Cl, forming an ionic bond (see Figure 2.8).

In the formation of ionic bonds, the transfer of electrons is such a common process that we define two terms that apply specifically to this type of bond:

Oxidation: loss of electrons

Reduction: gain of electrons

Thus, in the formation of NaCl, the neutral Na atom is oxidized into a positive Na^+ ion when it loses an electron. A neutral Cl atom is reduced into a negative Cl^- ion by gaining an electron.

TABLE 2.1 Group Number Characteristics of Commonly Used Chemical Elements in Wafer Fabrication

Group Number	Characteristics
IA	• 1 valence electron that is easily given up; low electronegativity • Highly unstable • Very reactive; explosive • Form ionic bonds • Prefer not to use the metals in this group due to contamination issues
IIA	• 2 valence electrons • Somewhat unstable • Quite reactive • Prefer not to use metals in this group
IIIA	• 3 valence electrons • Dopant elements (primarily B) added to semiconductor material • Common interconnect conductor material (Al)
IVA	• 4 valence electrons • Semiconductor materials • Form covalent bonds
VA	• 5 valence electrons • Dopant elements (primarily P and As) added to semiconductor material
VIA	• 6 valence electrons
VIIA	• 7 valence electrons; readily accept electrons; high electronegativity • Corrosive • Very reactive • Form ionic bonds • Useful in some semiconductor applications; used as etching and cleaning compounds
VIIIA	• 8 valence electrons • Stable; nonreactive • Inert gas • Safe to use in some aspects of semiconductor manufacturing
IB	• Best metal conductors • Cu is replacing Al as primary interconnect conductor material
IVB – VIB	• Refractory (high melting temperature) metals commonly used in semiconductor manufacturing to improve metallization (especially Ti, W, Mo, Ta, and Cr) • React well with silicon to form stable compound with good electrical characteristics

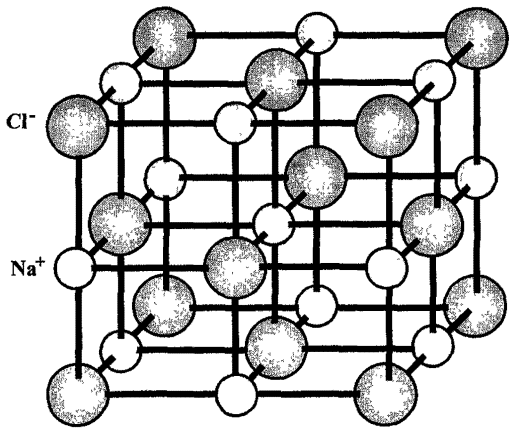


FIGURE 2.8 Ionic Bond for NaCl

Covalent Bonds

Another form of chemical bond is a covalent bond, where atoms of different elements share valence-shell electrons. The atoms share these valence electrons in order to attain a full valence shell and become stable. An example of a covalent bond is hydrogen chloride (HCl), where molecules consist of one hydrogen (H) atom forming a covalent bond with one chlorine (Cl) atom to form HCl (see Figure 2.9).

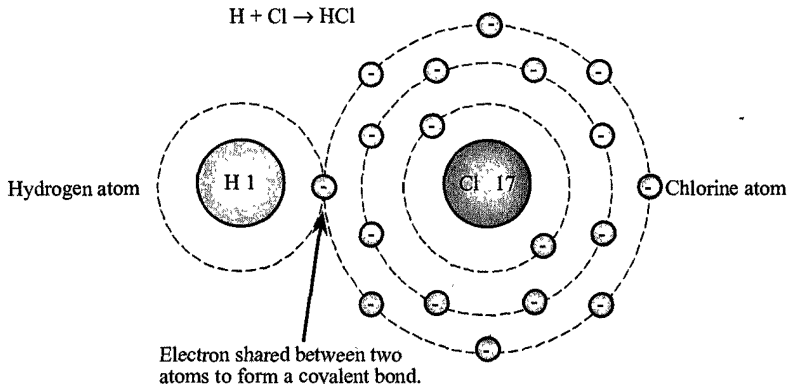


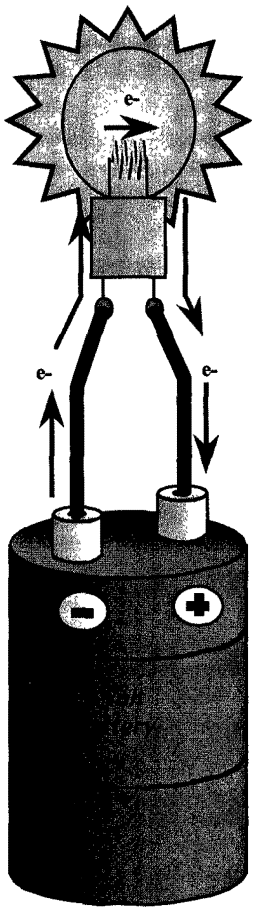
FIGURE 2.9 Covalent Bond for HCl

CLASSIFYING MATERIALS

Electrical current is the movement of charge carriers (electrons in metals, electrons or holes in semiconductors) from one point to another in a material under the influence of applied electric fields.[2] The unit for current is amps (see Figure 2.10). A major way of classifying electronic materials is based on how current flows through the material.

There are three different classes of materials based on the flow of electric current through these materials:

- Conductors
- Insulators
- Semiconductors



Copper wires provide connections that allow electrons to flow from the - terminal, through the filament inside the lamp, and back into the + terminal of the battery.

FIGURE 2.10 Electrical Current Flow
(Used with permission from Advanced Micro Devices)

Conductors

A material is a *conductor* if electrons readily flow through the material as electrical current. A good conductor has high current-carrying capability, referred to as current density. A conductor generally has few valence electrons in its outermost shell that are loosely bound and easily given up by the atom. Metals typically have this valence-shell configuration. Common conductor metals used in semiconductor manufacturing are aluminum, used for interconnect wiring between devices, or tungsten, used as the material for forming electrical interconnections between metal layers.

An example of a good metal conductor is copper (Cu), which recently has been introduced as a replacement for aluminum metallization for interconnecting the different devices on a microchip during wafer fabrication. Cu has twenty-nine electrons, with one electron in the valence shell that is relatively far from the nucleus. This electron is easily freed from the atom and is available to conduct electricity (see Figure 2.11).

Copper atom

Shell #	Maximum # e⁻ per shell	Actual # e⁻ per shell
K	2	2
L	8	8
M	18	18
N	32	1
Total #	60	29

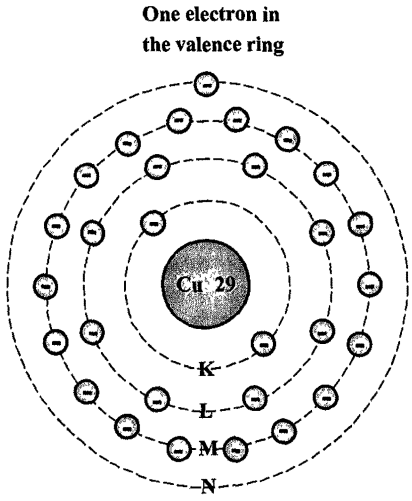


FIGURE 2.11 Flow of Free Electrons in Copper

Conductors require little energy to move a valence electron from its valence band into the conduction band. This electron movement leaves a gap in the covalent bond of the atom, which is referred to as a hole. For every conduction-band electron there must exist a valence-band hole. Valence electrons can jump into the position of the hole, which promotes electron movement and conduction. This condition is referred to as the *electron-hole pair.*

Very quickly after becoming a free electron in the conduction band, an electron gives up its energy and falls into a hole in the covalent bond of the valence band. This process is known as *recombination*. The time period from when an electron moves from the conduction band until recombination is known as the *lifetime* of the electron-hole pair. Thermal energy leads to the continual creation of electron-hole pairs followed by their recombination.

Conductivity and Resistivity ■ The property of a material to conduct electricity is its *conductivity*. Materials are usually characterized by their *resistivity* (ρ), or resistance to current flow. A material with lower resistivity has better conductivity. These two properties of resistivity and conductivity depend only on the material, not its geometry, and are related by the following formula:

$$\rho = \frac{1}{C}$$

where, ρ = resistivity in ohms-centimeter (Ω-cm)
C = conductivity

Resistivity of materials is important in semiconductor products because of its effects on the electrical operation of an integrated circuit. Resistivity can be deliberately designed into an IC in the form of an electronic component called the resistor. In this case the resistor serves as a desired control for current in a particular part of an electronic circuit. In some cases, resistivity is an

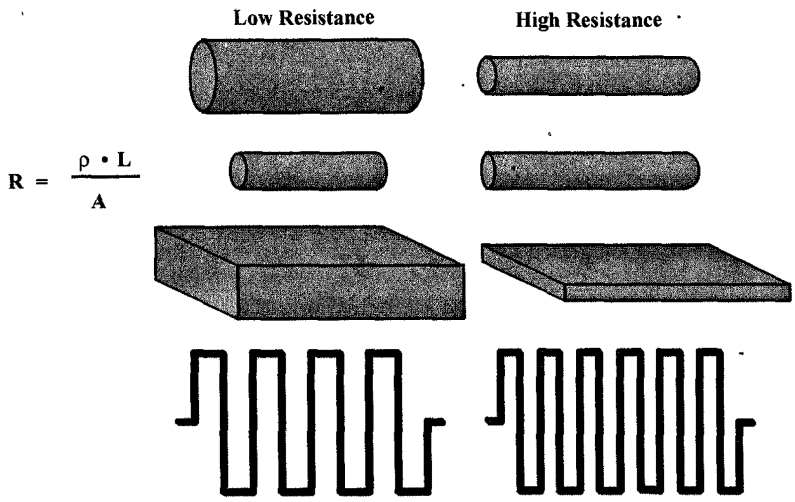


FIGURE 2.12 How Wiring Sizes Affect Resistance

undesirable characteristic of materials that introduce too much resistance to the flow of current, resulting in power losses and overheating within the IC.

Resistance. *Resistance* is the opposition to current flow and is accompanied by the dissipation of heat. A material with high resistance has high opposition to current flow. Resistance depends on both resistivity and the physical geometry of the material and is calculated by:

$$R = \frac{\rho l}{\text{area}}$$

where, R = resistance of the conductor material, in ohms
ρ = resistivity of the conductor material, in ohms-cm
l = length of the conductor, in cm
area = cross-sectional area of conductor = width $\times$ thickness

Reduced feature sizes in wafer fabrication makes resistance an important parameter. Smaller geometries cause increased resistance in interconnect lines, which is undesirable due to increased heat losses. Lower resistance is why copper is replacing aluminum as the primary interconnect wiring material. Another example of the application of resistance in semiconductor manufacturing is the sheet resistance measurement to control the thickness of conductive films (see Chapter 7). Since resistance is dependent on geometry, the sheet resistance of a thin film changes depending on the film's thickness.

Insulators

An insulator is a material that has a high resistance to current flow. Another term for an insulator is a *dielectric.* Insulators have no loosely bound electrons in the valence shell available for electrical conduction, with a high energy gap to keep valence band electrons out of the conduction band. Examples of insulators in everyday life are rubber, plastics, glass, and ceramics. Some insulator materials in semiconductor manufacturing are silicon dioxide (SiO_2), silicon nitride (Si_3N_4), and polyimide (a plastic material).

Purified and deionized (DI) water is a good example of an insulator, with a resistivity of about 18×10^6 ohms-cm, or 18 megohms-cm. There are insufficient free electrons in DI water to be able to sustain current flow from a small battery. However, the conductivity of an insulator can be improved by adding an impurity to it (see Figure 2.13). In the case of water, we can add common table salt. The salt dissociates in water into its basic ionic elements of sodium (Na^+) and chlorine (Cl^-), forming an electrolyte (a conducting solution). The net effect of these charged atoms is the same as having free electrons in a piece of copper wire—if there are sufficient charged atoms, current flow can be sustained. As we shall learn later, adding an impurity to alter an insulator material's conductivity is an important aspect of semiconductor technology.

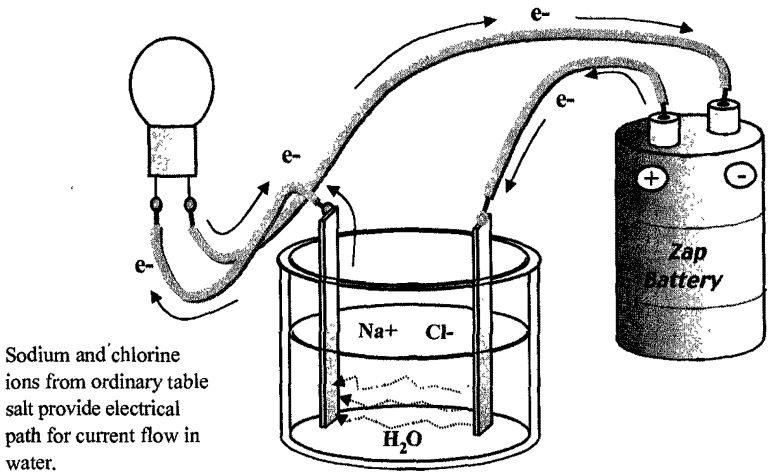


FIGURE 2.13 Adding an Impurity to Water to Improve its Conductivity

Capacitance ■ *Capacitance* is the storage of electrical charge on two conductive plates separated by a dielectric material (see Figure 2.14). The unit of measure for capacitance is farads and is usually expressed as picofarads (10^{-12} farads) for capacitance in integrated circuit structures. The amount of charge that can be stored by a capacitor varies depending on certain physical characteristics. These characteristics include the area of the plates, the spacing between the plates, and the quality of the insulation material between the plates, better known as the *dielectric constant, k* (farads/cm). The *k* value of air is 1, while that of glass is between 4 and 7.

We can use a simple circuit to explain how a capacitor holds an electrical charge (see Figure 2.15 on page 32). In this circuit, a 1.5 V battery is connected to a capacitor. When the switch is closed, electrons from the left metal plate are attracted to the positive (+) side of the battery. At the same time, electrons from the negative side (-) of the battery flow to the right plate to balance the deficiency of electrons from the left side. The initial current stops due to the dielectric material between the plates. The result, however, is a difference in potential between the two plates. An electrostatic field exists between the charged plates.

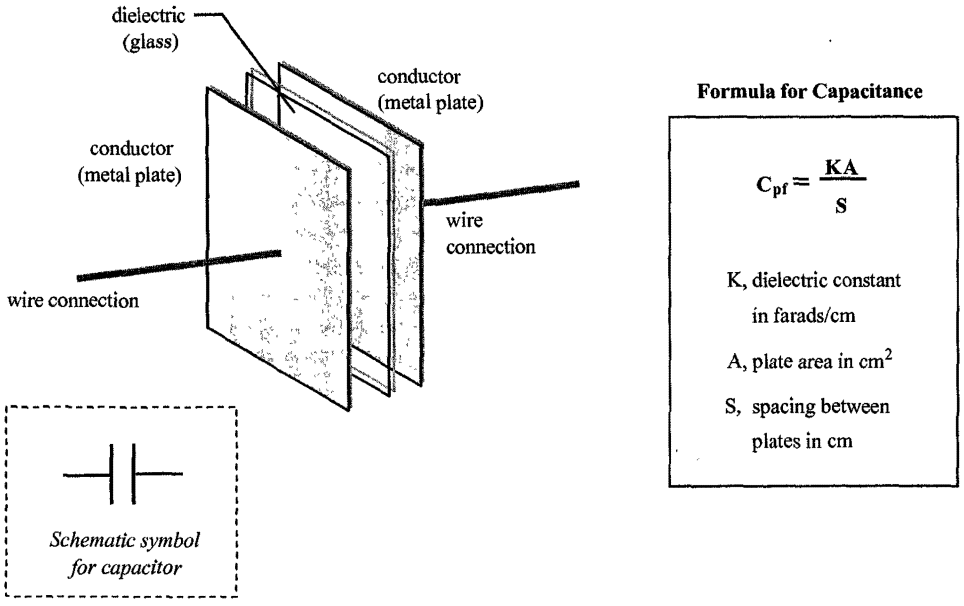


FIGURE 2.14 Basic Capacitor Structure
(Used with permission from Advanced Micro Devices)

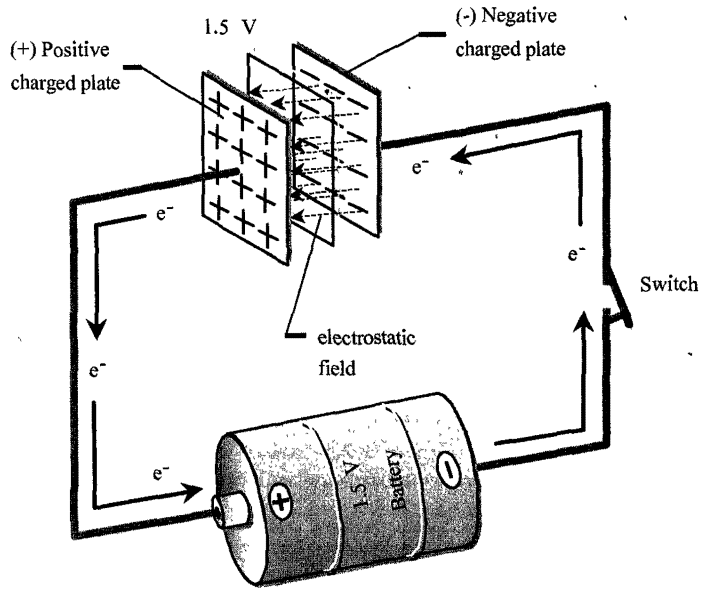


FIGURE 2.15 Battery Charges a Capacitor
(Used with permission from Advanced Micro Devices)

If the battery and circuit are removed from the capacitor, the left plate is positively charged and the right plate is negatively charged (see Figure 2.16). The overall difference in potential is equal to the original battery value and remains as long as there is no other way for electrons to flow to the opposite side of the capacitor (in reality, some leakage does occur through the dielectric). Understanding these concepts regarding how a capacitor works is crucial for understanding how a field-effect transistor (FET) works (see Chapter 3).

Dielectric Constant. The dielectric material is a key component of capacitors. This material can change the capacitance of a capacitor by more effectively concentrating the electric field occurring between the two conductors. This dielectric constant, k, has become an important parameter in semiconductor performance. As current flows through adjacent metal conductors that interconnect the devices on the chip (referred to as wiring), it is desirable to have a low-k dielectric material to minimize capacitance losses (see Figure 2.17).

Semiconductors

The third type of material is *semiconductors*. These materials are special because they can function as either conductors or insulators. A semiconductor material has a small energy gap (i.e., 1.11 eV for silicon) that is a value between an insulator (>2 eV) and a conductor. This energy gap permits the electrons to jump from the valence band into the conduction band when energy is supplied. This

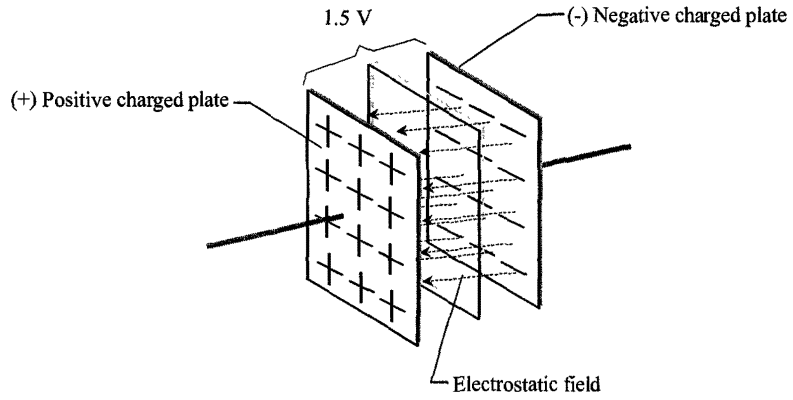
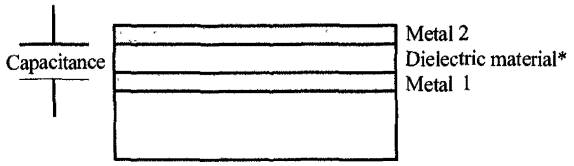


FIGURE 2.16 Capacitor Holds a Charge



** Low k dielectric reduces capacitance between metal layers.*

FIGURE 2.17 Low-*k* Dielectric Material

action occurs when a semiconductor is heated, thus improving conductivity with an increase in temperature (which is the opposite condition for a conductor).

The most important semiconductor material in wafer fabrication is silicon, used as the semiconductor wafer substrate in over 85% of all chips.[3] This text will only address silicon because of its predominant use in the semiconductor industry. A short review of some of the other semiconductor materials and their applications is given at the end of this chapter.

SILICON

Silicon is an elemental semiconductor material because it has four valence-shell electrons, along with the other elements located in Group IVA of the periodic table (see Figure 2.18 on page 34). The number of valence-shell electrons in silicon places it midway between the valence-shell condition of a good conductor (one valence electron) and an insulator (eight valence electrons).

Silicon is not found pure in nature. It must be refined and purified to become pure for semiconductor manufacturing (see Chapter 4). It usually occurs as silica (silicon dioxide, or SiO_2) and other silicates. Silica is sand and is a primary ingredient of glass. Other forms of SiO_2 are rock crystal, quartz, agate, and opal.

The melting point of silicon is 1412°C. Silicon is a hard and brittle material that fractures easily if deformed, much like a piece of glass. It can be polished to a mirrorlike finish. Silicon exhibits many of the same characteristics as metals and at the same time exhibits characteristics of nonmetals. This is why silicon is classified as a semiconductor, midway between the conductors (metals) and insulators (nonmetals) on the periodic table.

Pure Silicon

Pure silicon is referred to as *intrinsic silicon,* with no contaminants or impurities from other substances. The silicon atoms in pure silicon bond together through covalent bonds to share electrons and complete their valence shells (see Figure 2.19 on page 34).

Many of silicon's properties result from this strong covalent bonding. The covalent bonds in pure silicon hold the atoms together to form a solid, electrically stable material that is an insulator. Pure silicon is a poor conductor because all valence shells are fully occupied in covalent bonds. In this pure form, silicon is not useful as a semiconductor.

The solid material formed when two or more silicon atoms bond together in this set, repeatable pattern, as the one shown in Figure 2.19 on page 34, is referred to as a *crystal.* A crystal is a smooth, glassy solid that forms a lattice structure with a three-dimensional shape. An example of a crystal material is window glass. We will study more about crystals in Chapter 4 when we discuss silicon wafer preparation.

Why Silicon?

Germanium was the first material used for semiconductors in the 1940s and early 1950s, but as we learned in Chapter 1, it was soon replaced by silicon. Why was silicon chosen as the predominant semiconductor material? There are four main reasons for the selection of silicon as the primary semiconductor material:

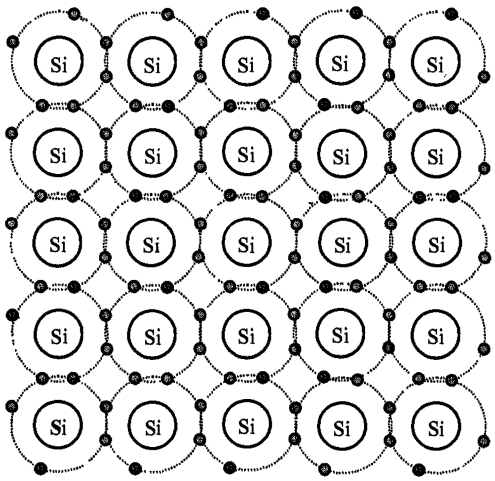
- Abundance of silicon
- Higher melting temperature for wider processing range

- Wider temperature range of operation
- Natural growth of silicon dioxide

Semiconductors

Group IVA	
C, Carbon	6
Si, Silicon	14
Ge, Germanium	32
Sn, Tin	50
Pb, Lead	82

FIGURE 2.18 Group IVA Elemental Semiconductors



Silicon atoms share valence electrons to form insulator-like bonds.

FIGURE 2.19 Covalent Bonding of Pure Silicon

Silicon is the second most abundant element on earth and makes up about 25% of the earth's crust. If processed properly, silicon can be refined into ample quantities of the very pure form necessary for semiconductor fabrication which leads to lower costs. Silicon's melting temperature of 1412°C is much higher than the melting temperature for germanium which is 937°C. This higher melting temperature permits silicon to withstand high-temperature processing. Another advantage to using silicon is that a semiconductor device made from silicon can function over a wider temperature range than germanium, which increases the semiconductor's applications and reliability.

Finally, an important reason for using silicon as a semiconductor material is the ability to naturally grow *silicon dioxide* (SiO_2) on the silicon surface (see Figure 2.20). SiO_2 is a high-quality, stable electrical insulator material that also serves as a good chemical barrier to protect silicon from external contaminants.[4] Electrical stability is important to avoid leakage between adjacent conductors in an IC. The ability to grow stable, thin SiO_2 material is fundamental to the fabrication of high-performance metal-oxide semiconductor (MOS) devices. SiO_2 has mechanical properties similar to silicon, which allows high temperature processing without excessive wafer warpage.

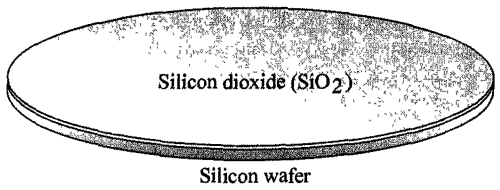


FIGURE 2.20 SiO_2 on Silicon Wafer

Doped Silicon

Silicon in its pure state has little use in semiconductor technology. However, the structure of silicon can be altered to greatly enhance its conductivity by adding small amounts of other elements to the material through a process known as doping. *Doping* is the process of adding certain elements to pure silicon to improve the conductivity of the semiconductor (see Figure 2.21). For example, the electrical resistivity (ρ) of pure silicon is approximately 2.5×10^5 ohms-cm. If only one in

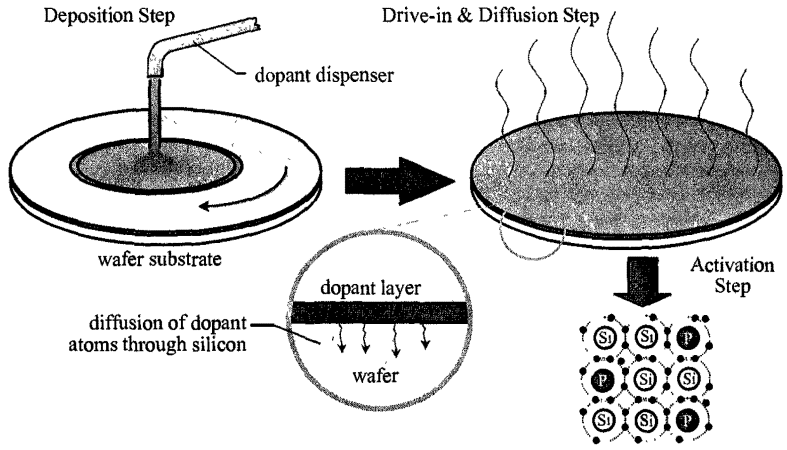


FIGURE 2.21 Doping of Silicon
(Used with permission from Advanced Micro Devices)

every one million silicon atoms is replaced by one atom of arsenic, resistivity will drop to 0.2 ohms-cm.[5] This is an improvement in conductivity of 1,250,000 times.

The elements added during doping are referred to as *dopants* or *impurities* because the silicon is no longer pure. In other words, we dope the silicon with an impurity so that it will conduct electricity. The more impurity added then the higher the conductivity (or the lower the resistivity). Note that we use the term impurity to indicate that we have added another element to the silicon. We intentionally add the impurity to increase the conductivity of silicon. Doped silicon is also known as *extrinsic silicon.*

The concept of doping the silicon with an impurity to improve electrical conductivity is a critical aspect of semiconductor fabrication. If we can introduce impurities that alter the silicon's resistivity and then control when the silicon performs as an insulator or as a conductor, then we have the essence of solid-state technology.

Dopant Materials ■ Silicon is located in Group IVA of the periodic table and has four valence electrons. Elements from the two adjacent groups are commonly used for doping: Group IIIA and Group VA (see Figure 2.22). Group IIIA elements are referred to as *trivalent* because of three valence electrons, whereas elements from group VA are *pentavalent* because of five valence electrons. The trivalent dopant increases the number of holes (positive doping or p-type), whereas the pentavalent will increase the number of free electrons (negative doping or n-type).

Acceptor Impurities		Semiconductor		Donor Impurities	
Group III (p-type)		**Group IV**		**Group V (n-type)**	
Boron	**5**	Carbon	6	Nitrogen	7
Aluminum	13	**Silicon**	**14**	**Phosphorus**	**15**
Gallium	31	Germanium	32	**Arsenic**	**33**
Indium	49	Tin	50	**Antimony**	**51**

** Elements underlined are the most commonly used in silicon-based IC manufacturing.*

FIGURE 2.22 Silicon Dopants

When trivalent dopant atoms are added to silicon, the resulting material is called a *p-type silicon.* The trivalent dopants are referred to as *acceptors* (they accept an extra mobile electron), with the most common acceptor element being boron. When a pentavalent element is added to pure silicon, the resulting material is referred to as an *n-type silicon.* Pentavalent dopants are referred to as

donors (they donate an extra mobile electron), and are typically either phosphorous, arsenic, or antimony. The different doping elements are listed in Table 2.2.

TABLE 2.2 Dopant Elements Commonly Used in Semiconductor Manufacturing

Trivalent Dopants (p-type, positive doping, acceptor)	Pentavalent Dopants (n-type, negative doping, donor)
Boron (B)	Phosphorus (P)
	Arsenic (As)
	Antimony (Sb)

N-Type Silicon. For an n-type silicon, there are more electrons than valence-band holes. This is shown in Figure 2.23 for silicon with a pentavalent phosphorous dopant atom.

The silicon atoms will form covalent bonds with the donor phosphorous atom, each sharing one phosphorous electron. However, the fifth phosphorous electron is not bound to any of the surrounding silicon atoms. Because of this, the fifth phosphorous electron needs little energy to break away and enter the conduction band. For n-type silicon, conduction-band free electrons are the majority carriers, of which the material has an abundance. There are also many fewer minority carrier valence-band holes.

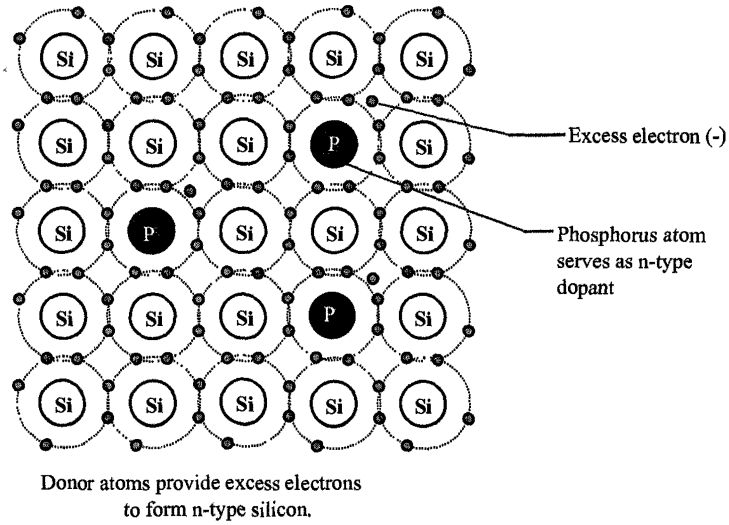


FIGURE 2.23 Electrons in n-Type Silicon with Phosphorous Dopant

One electron in the conduction band is not significant for conduction. However, when we dope silicon with a dopant, we add literally millions of dopant atoms, producing many electrons that are not part of the covalent bonds. There is a high amount of movement between electrons and holes. Negative electrons are attracted to the positive holes. Electrons enter the conduction band relatively easily. If a voltage is applied to this material, the electrons can be made to flow through the material as electric current (see Figure 2.24).

Note that the doped silicon is still electrically neutral (this is true for n-type or p-type silicon). In the case of n-type silicon, this is because each phosphorous atom still has the same number of protons as electrons, as do the silicon atoms. Thus the overall number of protons and electrons in the semiconductor is still equal and the result is a net charge of zero. What is not equal is that there are many more conduction-band electrons (majority carriers) than there are valence-band holes (minority carriers).

P-Type Silicon. In the p-type silicon shown in Figure 2.25, the boron atom is a p-type acceptor dopant that forms a covalent bond with four adjacent silicon atoms. The boron acceptor atoms provide a deficiency of electrons caused by the lack of a fourth electron in the boron

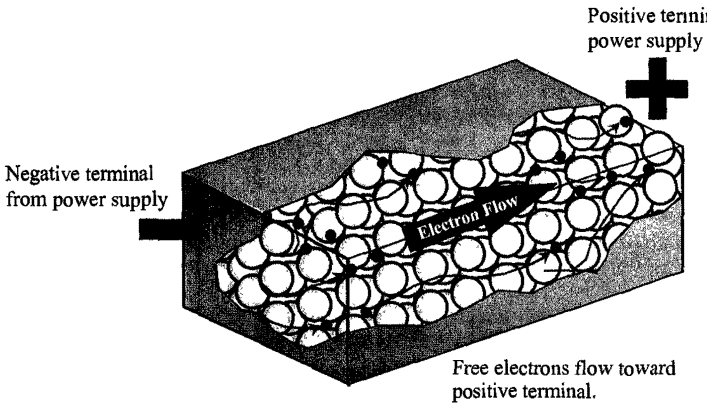


FIGURE 2.24 Flow of Free Electrons in n-Type Silicon

atom, thus creating the p-type silicon. There is an excess of holes (absence of electrons) in the valence shell.

Since there are many more valence-band holes than conduction-band electrons, the holes are the major current carriers in p-type material. The holes are referred to as *majority carriers* and electrons are the *minority carriers.* If a dc voltage is applied across a p-type semiconductor, then the large number of holes in the material will attract electrons from the negative terminal of the voltage source into the p-type semiconductor. This is current flow in a p-type material (see Figure 2.26 on page 38). The holes appear to move because each time an electron moves into a hole it serves to create a hole in its previous position. The holes seemingly move in the opposite direction as the electrons.

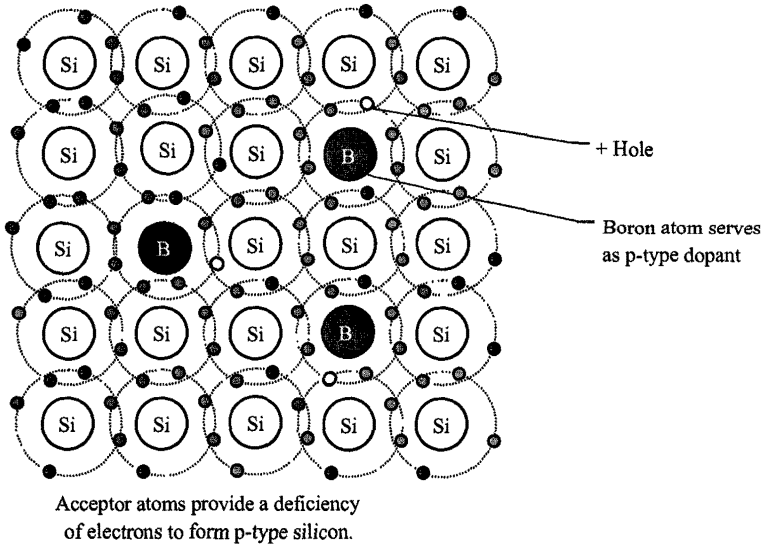


FIGURE 2.25 Holes in p-Type Silicon with Boron Dopant

Resistivity of Doped Silicon. We obtain precise control of silicon's resistivity by introducing dopants into the silicon crystal structure. The concentration of the dopant atom placed in the silicon crystal defines how well the material will conduct electrical current. Pure silicon has a resistivity of about 250,000 Ω-cm and is an insulator. Compare this to copper, a good conductor, with a resistivity of 1.7 μΩ-cm (0.0000017 Ω-cm). By adding the proper type and concentration of dopants to pure silicon, the doped silicon's resistivity is decreased and conductivity is improved (see Figure 2.27 on page 38). For a given resistivity there is less concentration of n-type dopants than p-type dopants. This is because it takes less energy to move an electron than to move a hole.

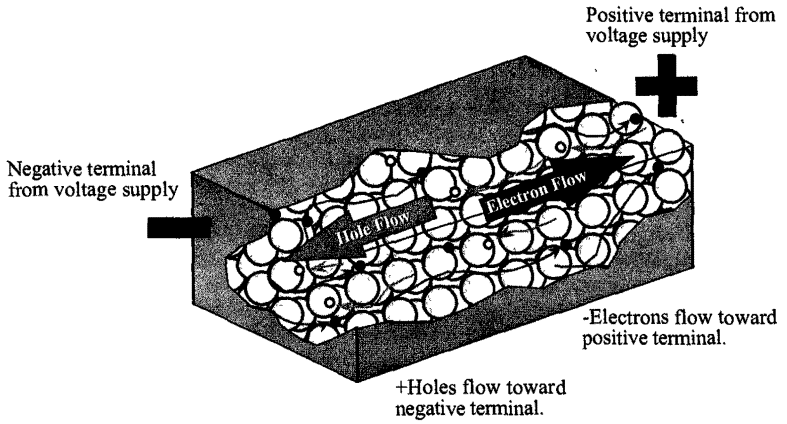


FIGURE 2.26 Flow of +Holes in p-Type Silicon

It only takes a small amount of dopant (from as little as 0.000001% to 0.1%) to make silicon a useful conductor. This is important for the fabrication of semiconductor devices on wafers. The amount of dopant in the silicon, or concentration, is carefully controlled during semiconductor fabrication to attain a precise resistivity. The ion implantation process for adding dopants to silicon will be covered in Chapter 17.

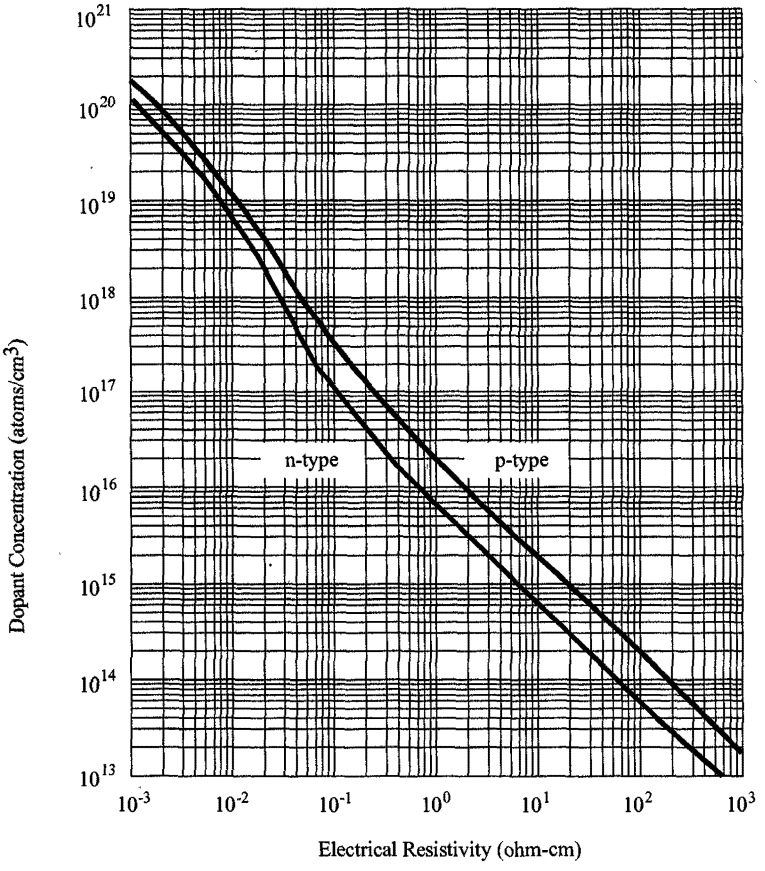


FIGURE 2.27 Silicon Resistivity Versus Dopant Concentration
Redrawn from S. Ghandi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide*, John Wiley & Sons, 1994.

pn Junctions

We dope pure silicon with pentavalent or trivalent elements to obtain either an n-type or p-type semiconductor. The type and concentration of the dopant element determines whether conduction occurs by hole or by electron and determines the final value of the silicon's resistivity. The number of carriers, and therefore the resistivity, is determined by the net donor and acceptor atom concentration in the silicon crystal. Because of this, an n-dopant can be implanted or diffused into a p-region to convert that region into an n-region (or vice versa). To achieve this, the concentration of the n-type donor dopant must exceed that of the p-type acceptor dopants in the region.

The ability to incorporate both n-type and p-type regions in the silicon crystal is beneficial because semiconductor devices require both of these regions in order to function as a semiconductor. It is the junction between the n-type and p-type regions that is important and creates the useful characteristics of silicon as a semiconductor. This junction is referred to as a *pn junction* (see Figure 2.28).

The pn junction is the essence of solid-state electronics and is the basis for how semiconductor wafers can achieve their unique ability to act as an insulator and a conductor, depending on the bias voltage applied to the junction. The specifics of how a pn junction functions as a semiconductor are discussed in Chapter 3. The creation of a pn junction is almost universally done in wafer fabrication with ion implantation (see Chapter 17).

Remember that a pn junction is formed between two pieces of essentially the same material. The p-type and the n-type materials scarcely differ except for a minute amount of a dopant. The n-type material has extra mobile electrons from the donor impurity, whereas the p-type material has extra mobile holes. It is deceptive to speak of one material meeting the other.[6] The junction is so intimate that the n-type and p-type materials are formed from one continuous solid. The silicon crystal with a pn junction continues to look like and in most ways behave like the solid crystal material that it is.

The depth and definition of this junction is critical in semiconductor manufacturing. As device critical dimensions are reduced, the ability to precisely control both the pn junction in silicon (e.g., junction depth) and the concentration of dopants is becoming a major challenge for semiconductor chip fabrication.

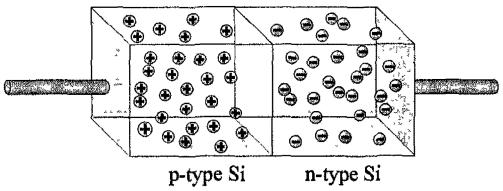


FIGURE 2.28 Cross Section of Planar pn Junction

ALTERNATIVE SEMICONDUCTOR MATERIALS

Germanium and silicon are the two elemental semiconductor materials from Group IVA with four valence electrons. We have learned that germanium was the first semiconductor material used for transistor fabrication and was replaced by silicon in the 1950s for process and performance reasons.

There are alternative semiconductor materials used for specific market applications. These are primarily the *compound semiconductors.* A major class of compound semiconductors are formed from Group IIIA and Group VA of the periodic table (often referred to as III-V compounds). An example is gallium arsenide (GaAs). In addition, other semiconductor compounds come from elements found in Groups IIA and VIA, and are referred to as II-VI compounds.

The need for alternative semiconductor materials is driven by the growth of markets that require IC performance beyond that capable by silicon semiconductors. The key IC performance factor is speed. Wireless and high-speed digital communications, space applications, and consumer markets such as the automotive industry, are developing special niche semiconductor markets that

are centered on high speed that can support higher signal frequencies. In addition, III-V compound semiconductors are used for light-emitting diodes (LEDs).

GaAs is the most common compound III-V semiconductor and will be discussed in the next section. There are variations of this compound, such as gallium nitride (GaN) for use in the fabrication of blue semiconductor lasers and LEDs. Another compound semiconductor is silicon germanium (SiGe), which has been researched for more than two decades and may compete against GaAs for market applications.

There is also some growth in II-VI semiconductors, with their market revenue expected to be $179 million in 2001.[7] The two principal II-VI materials are cadmium telluride (CdTe) and zinc sellenide (ZnSe). CdTe semiconductors are used primarily in infrared (IR) detection systems. ZnSe is another II-VI compound material used to fabricate blue LEDs.

Gallium Arsenide (GaAs)

Gallium arsenide (GaAs) is the most common III-V compound semiconductor material and is formed by combining gallium (Ga) from Group IIIA with the element arsenic (As) from Group VA. GaAs has greater electron mobility than silicon, so the majority carriers move faster than in silicon. There are also some attributes of GaAs semiconductor material that reduce parasitic capacitance (see Chapter 3) and signal losses. These result in ICs that are faster than those made with silicon.[8] The improved signal speed of GaAs devices permits them to react to high-frequency microwave signals and accurately convert them into electrical signals for communication systems. Semiconductors based on silicon are too slow to react to the microwave frequency. For these reasons, products for wireless and high-speed digital communications and high-speed optoelectronics devices are made from GaAs and other compound semiconductors.

An advantage of GaAs is the increased resistivity of the material, with values up to 10^8 Ω-cm. This makes it easier to isolate semiconductor devices fabricated in the GaAs substrate without loss in electrical performance. GaAs devices also exhibit a higher radiation hardness than silicon, which makes this material attractive for military and space applications.

A summary of some different semiconductor material properties is shown in Table 2.3. The properties of SiO_2 have been included for comparison.

TABLE 2.3 Comparison of Some Physical Properties for Semiconductor Materials

Property	Si	Ge	GaAs	SiO_2
Melting point (°C)	1412	937	1238	1700 (approx.)
Atomic Weight	28.09	72.60	144.63	60.08
Atomic Density (atoms/cm^3)	4.99×10^{22}	4.42×10^{22}	2.21×10^{22}	2.3×10^{22}
Energy Band Gap (eV)	1.11	0.67	1.40	8 (approx.)

A. Grove, *Physics and Technology of Semiconductor Devices,* John Wiley and Sons, 1967.

The primary disadvantage of GaAs semiconductor material is the lack of a natural oxide. This feature hinders the development of standard MOS devices that require the ability to grow a surface dielectric. Another problem with GaAs is the fragility of the material, which makes wafer handling a major issue in wafer fabrication. The cost of GaAs is as much as ten times higher than silicon because of the relative scarcity of gallium and the energy used in its purification process. A final point to note is that the extreme toxicity of arsenic requires special controls for the equipment, process, and waste disposal facilities. These precautions have a significant effect on the manufacturing cost of GaAs semiconductors.

SUMMARY

All matter is made up of atoms. The number of valence-shell electrons in an atom is one of several factors that influences its chemical and physical properties and its ability to bond with other atoms. The energy-band theory of solids explains how electrons change orbital levels between the valence band and conduction band to be an insulator, conductor or semiconductor. Chemical elements are described in the periodic table, which is used to explain how ionic and covalent bonds form and how certain elements are used in semiconductor manufacturing. All material is classified as either a conductor, which conducts current, an insulator, which does not pass current, or semiconductor, which conducts current under certain conditions. Resistivity is a material property, while resistance depends on both resistivity and the geometry of the material. A capacitor uses a dielectric to store an electric charge.

Silicon is the most common semiconductor material because of its abundance, high melting temperature, wide temperature range, and ability to oxidize. Oxidation permits masking of the silicon surface and the ability to dope silicon with impurities to make it conductive. Silicon is doped with trivalent (p-type acceptor dopant) and pentavalent (n-type donor dopant) elements. Introducing a specific concentration of dopants into the silicon crystal permits the precise control of silicon's resistivity. When p-type and n-type silicon form a junction in the silicon crystal, then silicon can act as either an insulator or a semiconductor. There are alternative semiconductor materials, primarily the compound semiconductors, of which gallium arsenide is the most common.

KEY TERMS

element
molecule
compound
valence shell
insulators
conductors
semiconductors
cation
anion
ionization
atomic mass unit (amu)
electron-hole pair
recombination
lifetime
conductivity
resistivity
resistance
capacitance
silicon
intrinsic silicon
crystal
silicon dioxide
doping
dopants (or impurities)
extrinsic silicon
trivalent
pentavalent
p-type silicon
acceptors
n-type silicon
donors
majority carriers
minority carriers
pn junction
compound semiconductors
gallium arsenide

REVIEW QUESTIONS

1. What is matter?
2. Describe the model for atomic structure, including the proton, neutron, and electron, and state the charge of each.
3. Describe an element, a molecule and a compound.
4. What is a valence electron and why are valence electrons important for forming compounds?
5. How does energy-band theory explain the difference among insulators, conductors, and semiconductors? How does the energy gap change for each material?
6. How is an ion formed? What is a cation and an anion?
7. What does the group number on the periodic table represent in terms of valence electrons? What is the difference between the atomic number and the atomic mass number?
8. Define the atomic mass unit (amu).
9. How is an ionic bond formed? How is a covalent bond formed? What is oxidation and reduction and how are these processes related to the formation of a bond?
10. State the three classes of materials and explain how electric current flows in each one.
11. What is resistivity? What is conductivity? How is resistivity related to conductivity?
12. What is resistance? State the resistance formula.
13. What is another term for an insulator material?
14. What is capacitance? What is the dielectric constant and why is this concept important to semiconductor manufacturing?
15. Name the most common semiconductor material and give four reasons why it is so commonly used.
16. What is the difference between intrinsic and extrinsic silicon?
17. Describe silicon dioxide. Discuss why the oxidation of silicon is important.
18. What is doping? Why is doping important for semiconductor silicon?
19. Give one example of a trivalent doping element and three examples of a pentavalent dopant element. From what group number of the periodic table is each of these elements?
20. Describe the n-type doping of silicon with phosphorus.
21. n-type silicon has what type of majority and minority carriers? Give the majority and minority carriers for p-type silicon.
22. What characteristics determine how well silicon will conduct electricity?
23. Describe a pn junction and explain why it is important for a semiconductor.
24. Compound semiconductors come from what group numbers of the periodic table?
25. What is the advantage of gallium arsenide over silicon?
26. What is the primary disadvantage of gallium arsenide over silicon?

REFERENCES

1. Basic chemistry concepts referenced from M. Silberberg, *Chemistry, The Molecular Nature of Matter and Change,* (St. Louis: Mosby, 1996).
2. S. Wolf and R. Tauber, *Silicon Processing for the VLSI Era, Vol. 1—Process Technology* (Sunset Beach: Lattice Press, 1986), p. 118.
3. *Solid State Technology* (January 1998): p. 18.
4. F. Seitz and N. Einspruch, *Electronic Genie: The Tangled History of Silicon,* Urbana, (University of Illinois Press, 1998), p. 175.
5. G. Anner, *Planar Processing Primer,* (New York: Van Nostrand Reinhold, 1990), p. 57.
6. A. Holden, *Conductors and Semiconductors,* (Bell Telephone Laboratories, 1964), p. 109.
7. "Growth Seen for II-VI Materials," *Semiconductor International* (March 1997): p. 57.
8. S. Ghandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide,* 2nd ed., (New York: Wiley, 1994), p. 2.

CHAPTER 3

DEVICE TECHNOLOGIES

There are many different types of semiconductor devices, each meeting a functional need for users. This chapter introduces various electronic components and explains how they are constructed in silicon for use in the evolution of specific IC technologies.

OBJECTIVES

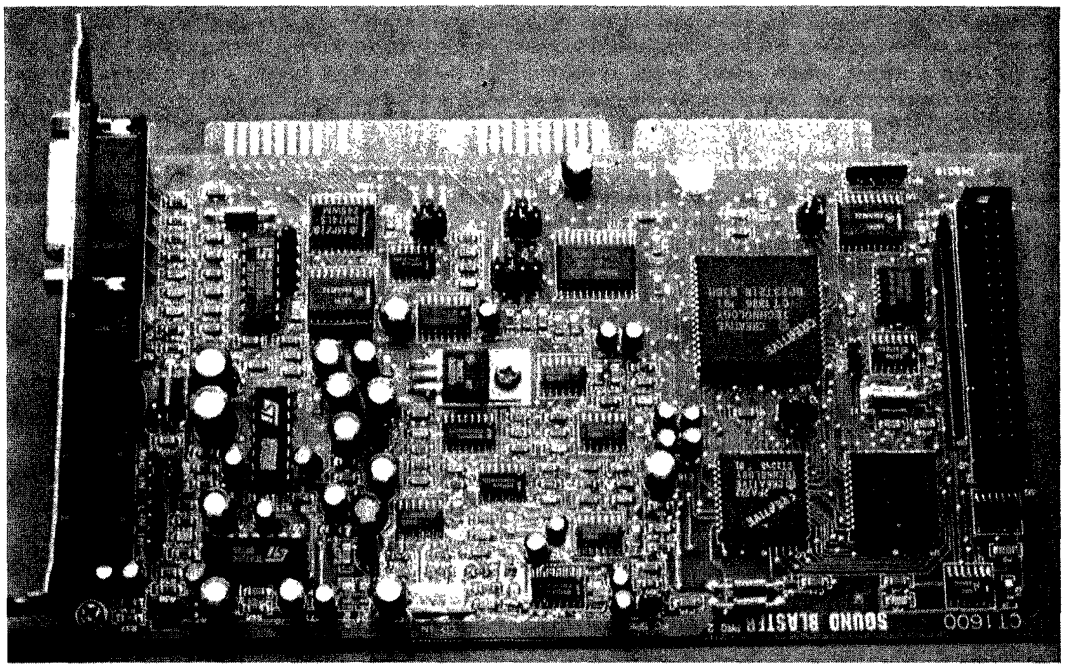
After studying the material in this chapter, you will be able to:

1. Identify differences between analog and digital devices and passive and active components. Explain the effects of parasitic structures in passive components.
2. Describe the pn junction, discuss why it is important, and explain reverse and forward biasing.
3. State the characteristics of bipolar technology and the bipolar junction transistor in terms of function, biasing, structure, and applications.
4. Explain the basic characteristics of CMOS technology, including the field-effect transistor, biasing, and the CMOS inverter.
5. Explain the difference between enhancement mode and depletion mode MOSFETs.
6. Explain the effects of parasitic transistors and the implications for CMOS latchup.
7. Give examples of IC products and state some applications of each.

INTRODUCTION

Electronic devices used in microchips are constructed in the silicon substrate. Common microchip devices include resistors, capacitors, fuses, diodes, and transistors. Their integration on a silicon substrate is the basis for IC wafer fabrication technology.

How a particular electronic device is formed in silicon is referred to as a *structure.* There are literally thousands of different semiconductor device structures and we will evaluate only a few of them. In this chapter, we discuss the actual formation of the devices in order to understand how they can perform in their many applications. Also covered in this chapter is a review of the various classifications of IC products.



Components on PC Motherboard

CIRCUIT TYPES

Circuits constructed from electronic components can be classified into two basic types: analog and digital circuits. Each circuit type has applications where it is most beneficial, as illustrated on page 30.

Analog Circuits

In electronics technology an *analog circuit* is one in which the electrical data varies continuously over a range of voltage, current, or power values. Analog circuits can be designed to operate with direct current (DC), alternating current (AC), or a combination of the two, pulsating DC. Some examples of electronics products that operate primarily as analog (also known as *linear*) circuits are radio transmitters and receivers, audio recording and playback systems, X-ray machines, and automotive ignition systems. The magnitude of the input and output signals, however, may not always be set at any predetermined level. For example, when scanning radio stations on an AM/FM radio receiver, not all radio signals have the same signal strength. As a result, the volume control may have to be adjusted according to the strength of the incoming signal. A volume control on a radio, a manual thermostat on a wall heater, the light control for a household light fixture—these are all examples of common analog devices.

Digital Circuits

Digital circuits have operating signals that vary about two distinct voltage levels—a high and a low. A logic high represents a binary bit = 1 and a logic low represents a binary bit = 0. Digital circuits are associated with *logic devices* such as computers and calculators. Other applications of digital logic devices include clocks, handheld computer games, and barcode readers. Digital devices can be used to measure and/or control events that require either an on/off type of command or that can be controlled in discrete incremental changes that approximate the actions of a linear circuit. That is why today it is sometimes difficult to tell the difference between an analog device and a digital system. Exactly what voltage levels a high and a low are depends on a specific *device technology*. Here are two examples of these logic voltage levels:

Logic Family	High State = 1	Low State = 0
TTL	5 VDC	0.0 VDC
CMOS	3.5 VDC	0.0 VDC

Note: VDC is volts DC.

PASSIVE COMPONENT STRUCTURES

You learned in Chapter 2 about the differences in materials categorized as insulators, conductors, or semiconductors. Now you will begin to see how these basic materials are used in the construction of some very basic electronic components—for example, a resistor and a capacitor. These components are referred to as passive components for various reasons. *Passive components* can conduct electrical current regardless of how the component is connected across a voltage supply. A resistor, for example, will conduct current equally as well regardless of which end of the resistor is connected to the positive or negative terminal of a battery.

IC Resistor Structures

An IC resistor can be created out of metal films or from doped polysilicon or by diffusion of dopant impurities into specific areas of the substrate (see Figure 3.1).[1] The resistors are microstructures and, therefore, occupy very small areas of the substrate. The connection of the resistors to the chip circuitry is formed from conductor metals such as aluminum or tungsten in the form of contacts (see Chapter 12).

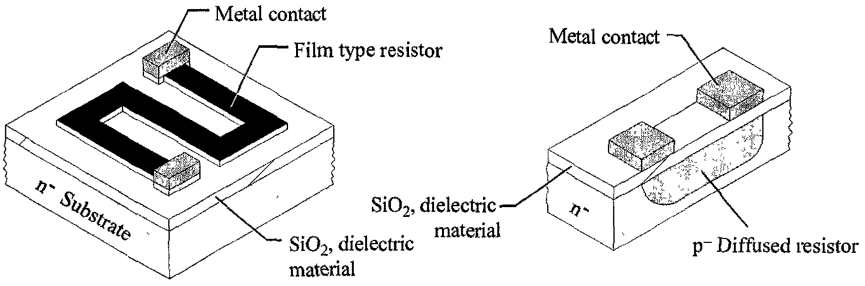


FIGURE 3.1 Examples of Resistor Structures in ICs

Parasitic Resistance Structures ■ *Parasitic resistance* is an unwanted resistance found in the design of components of an IC. It may exist in structures simply for reasons such as the size, shape, material type, dopant species, and quantity of dopant in the material. Parasitic resistance is an undesirable condition because it reduces the operational performance of IC devices. Figure 3.2 illustrates the locations of parasitic resistances in a transistor.

Parasitic resistances are cumulative, which means that the overall effect of these resistors in series is greater than a single resistance alone. The effect of parasitic resistances in IC devices becomes a detriment to the ability to reduce device feature sizes on chips. With higher circuit density comes higher resistances and overall degradation of electrical performance. Designers are aware of this problem and automatically calculate resistance losses into their design equations as well as select low-resistance metals for contacts and design process specifications to reduce the bulk resistance within active devices.

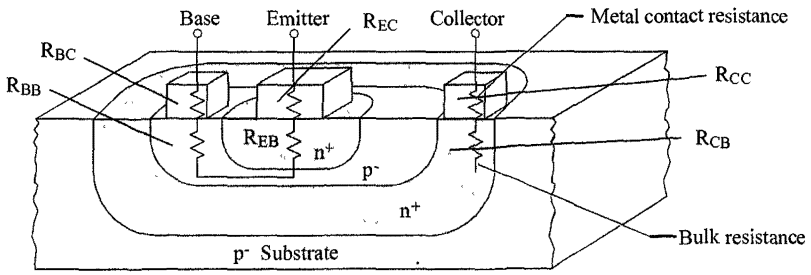


FIGURE 3.2 Cross Section of Parasitic Resistances in a Transistor

IC Capacitor Structures

You may recall in Chapter 2 that a simple capacitor is formed when two conductors are separated by a dielectric (insulator) material. In microchip manufacturing the dielectric material is usually silicon dioxide (glass, SiO_2), more commonly referred as simply *oxide. Planar capacitors,* which are built laterally on the substrate, may be formed from metallic films, doped polysilicon, or diffused areas of the silicon substrate. In general, capacitors are formed on substrates using four basic techniques (see Figure 3.3).

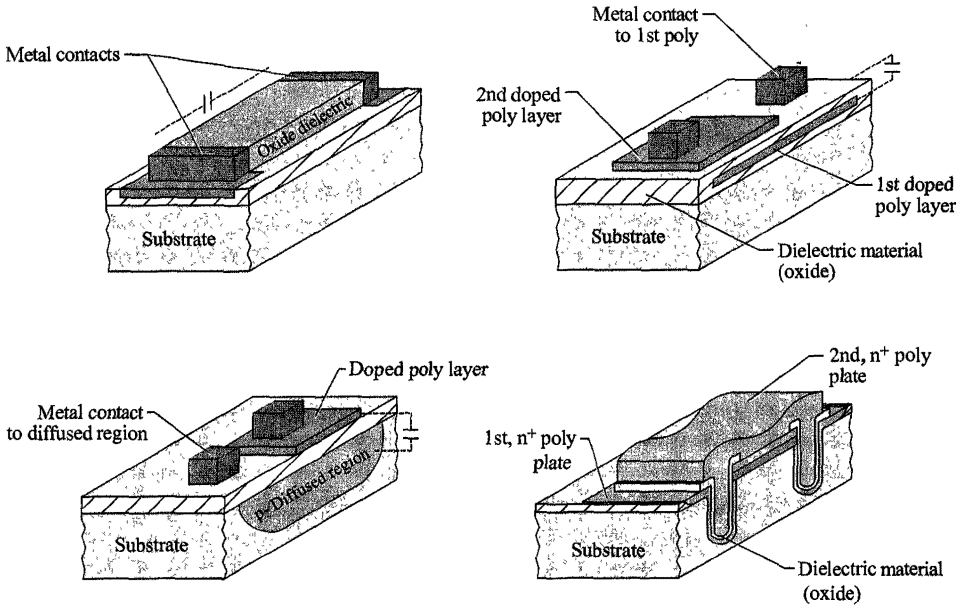


FIGURE 3.3 Examples of Capacitor Structures in ICs

Parasitic Capacitance Structures ■ Quite frequently capacitance occurs unintentionally as a result of the way materials are constructed on the substrate. This condition is referred to as *parasitic capacitance.*[2] An example of this would be two adjacent metal conductors with a dielectric material between them. This type of capacitance is undesirable for IC performance. In fact, parasitic capacitance interferes with the ability of electronic circuits to operate at faster switching speeds. In some cases parasitic capacitance may create instability in circuits, cause parasitic oscillations, and even create short-circuit paths for AC signals where they are not needed. Figure 3.4 illustrates the location of parasitic capacitance in between the electrodes of bipolar junction transistors and field-effect transistors.

ACTIVE COMPONENT STRUCTURES

Active components, such as diodes and transistors, offer some very different electronic control attributes that are not characteristic of passive components. They can be used to control the direction of current flow. Furthermore, active components can amplify small signals and can be used to create more complicated circuits, such as current and voltage regulators, oscillators, and logic gates. Active components require definite polarity (+ or -) when connecting these devices to power supplies. Active components utilize the flow of both electrons and holes in their operation.

The pn Junction Diode

A *pn junction diode* is formed any time there is a region of n-type semiconductor adjacent to a region of p-type semiconductor. A pn junction may be deliberately designed as a functional part of a larger integrated circuit or it may exist as a nonfunctioning diode in another part of the integrated

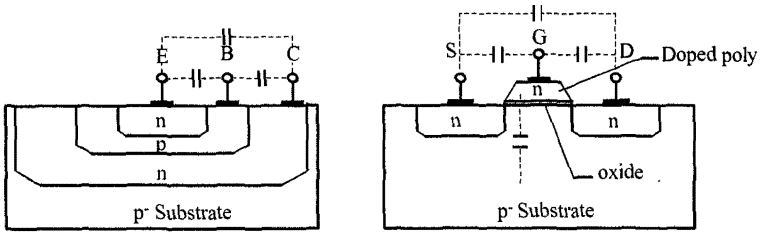


FIGURE 3.4 Examples of Parasitic Capacitance in Transistors

circuit. pn junctions are important for ICs because they are fundamental to the operation of both bipolar junction transistors and field-effect transistors.[3]

The pn junction diode is made in a single crystal of semiconductor material, e.g. silicon. As shown in Figure 3.5, one region of the substrate is doped heavily with donor dopant, such as arsenic, phosphorus, or antimony, to form an n-type silicon region. On the other side, acceptor impurity (e.g., boron) is used to create an p-type silicon region. The actual manufacturing procedures will be described later in Chapter 9. Typical metal contact materials used to connect the diode to the chip circuitry are aluminum, tungsten, titanium, or copper.

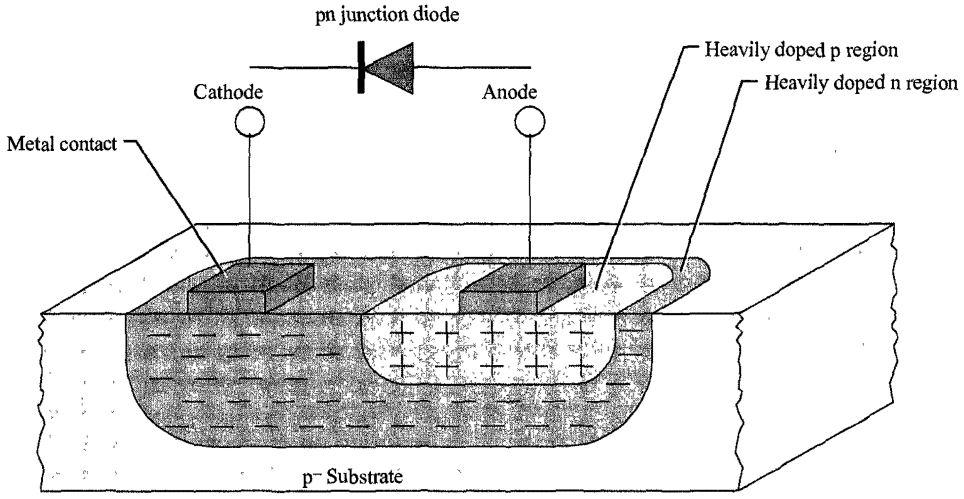


FIGURE 3.5 Basic Symbol and Structure of the pn Junction Diode

The open-circuit condition of the diode is shown in Figure 3.6 on page 48. Holes (+) make up the majority carriers in the p-type silicon while electrons (-) are the majority carriers in the n-type silicon. Initially, both sides of the diode are electrically neutral; however, some holes from the p region diffuse over to the n region, while electrons from the n region diffuse into the p region. These holes and electrons recombine at the junction of the diode. This recombination process has the net result of depleting each region of its majority current carriers. The p region has fewer holes than previously and the n region has been depleted of electrons. Thus, the term *carrier-depletion region* is used to describe the recombination area. The depletion of holes from the p region creates a negative charge in this area. Similarly, the loss of electrons from the n region results in a positive charge. The resulting charges from each region produce corresponding electric fields that oppose any further diffusion of carriers from the opposite side of the silicon crystal.

The net effect of the difference in charges across the depletion region is to create a potential difference across the junction. This results in a *barrier voltage* that has to be overcome before the diode can be operated.

Reverse Biasing the pn Junction Diode ■ Assume a voltage or bias supply (see the battery symbol in Figure 3.7) is connected across the pn junction as shown in Figure 3.7 on page 48. This bias

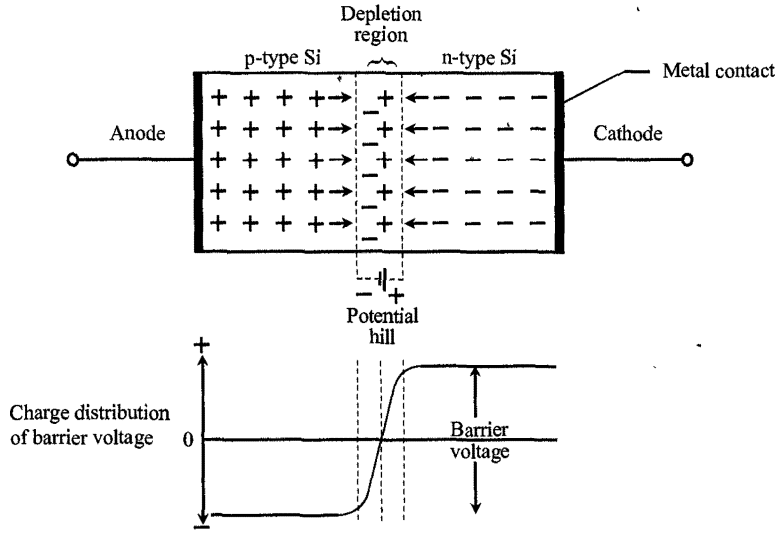


FIGURE 3.6 Open-Circuit Condition of a pn Junction Diode

configuration, called *reverse bias,* results in little or no conduction through the diode. The lamp is placed in the circuit as an indicator of current flow. Any appreciable current through the device can be attributed to minority current carriers in the diode.

Forward Biasing the pn Junction Diode ■ Figure 3.8 illustrates the effects of applying a *forward bias* to a pn junction diode. In this circuit electrons in the n region are repelled away from the negative terminal of the bias supply. Additional electrons are injected from the negative terminal to fill holes left vacant by the electrons in the n region. Similarly, holes in the p region are repelled away from the positive terminal of the bias supply. Holes are supplied from the positive side of the bias supply to balance the electrons supplied from the opposite side of the bias supply. Holes recombine with electrons at the junction and overcome the barrier voltage, greatly reducing its action as a barrier to current flow. Current flow will continue through the circuit as long as the bias supply can maintain a constant injection of holes and electrons through the diode.

Holes and electrons are drawn toward the junction to overcome the potential hill, thus, allowing current to flow through the circuit. Note that in Figure 3.8 solid arrows illustrate the direction of hole flow and broken arrows depict the direction of electron flow. Solid arrows also indicate conventional current flow.

The electrical characteristics of the diode are represented by the current-versus-voltage curve (see Figure 3.9). The graph in Figure 3.9 represents the forward and reverse bias characteristics of a typical silicon diode. Note the breakdown voltage at the knee of the forward bias curve. This voltage is characteristic of silicon diodes and is generally in the range of 0.6 to 0.8 volts. The reverse bias curve shows a point when the diode will conduct in the opposite direction when the junction voltage is exceeded.

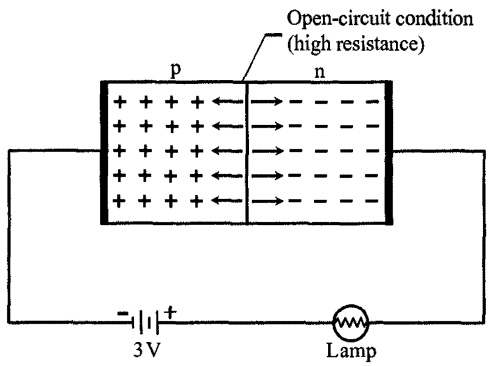


FIGURE 3.7 Reverse-Biased pn Junction Diode. Holes and electrons are drawn away from the pn junction, essentially creating an open-circuit condition.

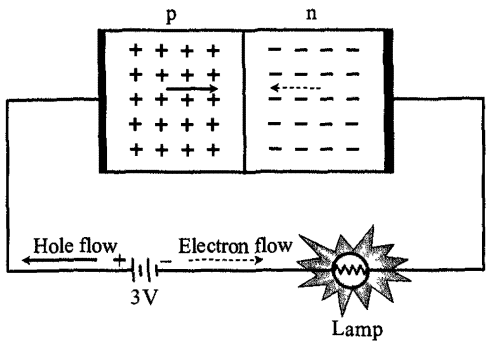


FIGURE 3.8 Forward-Biased pn Junction Diode

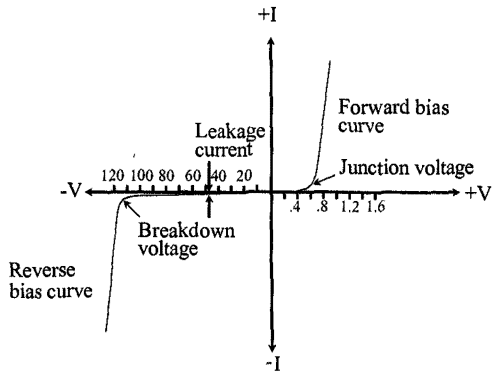


FIGURE 3.9 Forward and Reverse Electrical Characteristics of a Silicon Diode

The Bipolar Junction Transistor

The *bipolar junction transistor* (BJT) has three electrodes and two pn junctions. The entire transistor is constructed from a single semiconductor substrate. There are two variations of the bipolar junction transistor—*npn* and *pnp*. The simplified structures and schematic symbols of these transistors are shown in Figure 3.10. The electrodes are labeled *emitter, base,* and *collector*. The emitter arrows indicate the direction of hole flow through the respective transistor.

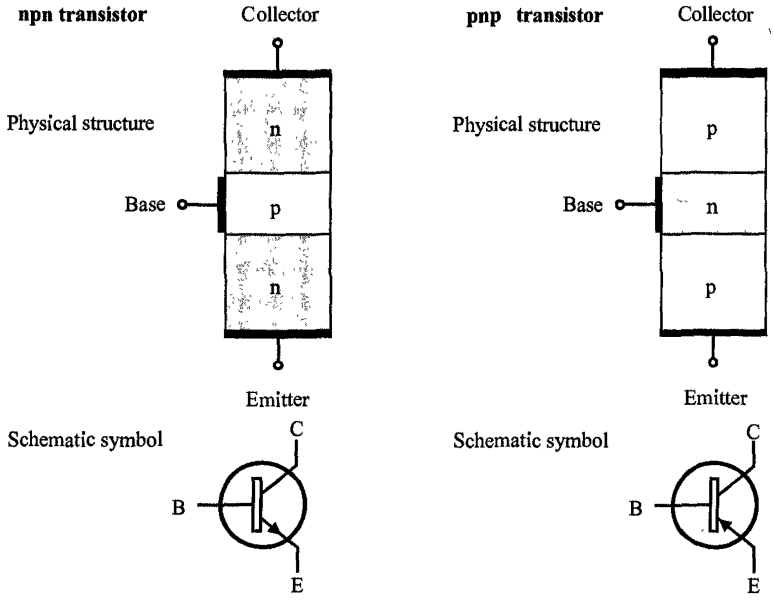


FIGURE 3.10 Two Types of Bipolar Transistors

Biasing the npn Transistor for Conduction Mode ■ The biasing scheme for proper operation of the npn transistor is shown in Figure 3.11 on page 50. Note the polarity of the two bias supplies. The emitter-base junction is forward-biased by a smaller bias supply, i.e., a simple 1.5-volt cell. The collector bias supply, a 3-volt battery, is connected in reverse polarity relative to the n-type collector. The emitter is shown as the reference point (ground) for all voltages. In this circuit configuration, the base serves as the input to the transistor and the collector is the output. The lamp serves the purpose as an output load and as an indicator of current flow through the collector.

There can be no conduction through the collector of the transistor without a completed circuit path through the emitter-base (E-B) junction. This nonconducting mode is shown in Figure 3.11 on page 50. As long as switch S_1 remains open, the pn junction of the emitter and base remains in a nonconducting mode (open circuit). This condition will not allow for current flow from emitter to

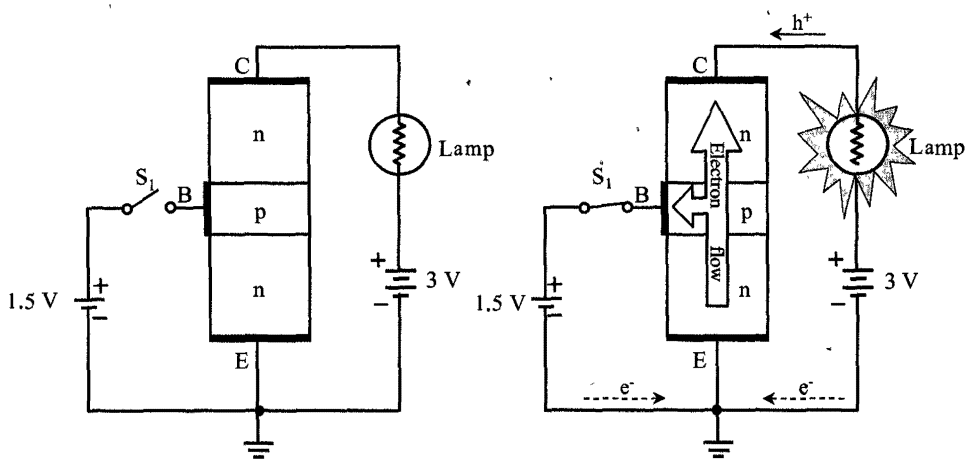


FIGURE 3.11 npn Transistor Biasing Circuit

collector. Closing switch S_1 forward-biases the E-B junction (the same as a diode). Conduction through the E-B junction immediately reduces the barrier and allows electrons to be injected from the negative side of the 3-volt battery through the E-B region and into the collector. At the same time holes are injected from the positive side of the collector battery through the lamp and then into the top of the collector to recombine with electrons. The lamp now indicates current flow through the output section of the transistor. This condition will continue until a change in S_1 or one of the circuit elements is disconnected.

Note that in Figure 3.11 solid arrows illustrate the direction of hole flow and broken arrows depict the direction of electron flow. Solid arrows also indicate conventional current flow.

Biasing the pnp Transistor for Conduction Mode ■ The operation of the pnp transistor is similar to the npn except that the bias supply connections are reversed. Conduction of a pnp transistor requires the same conditions as with the npn: first, the E-B junction must be forward biased (See Figure 3.12), and second, the collector supply voltage, V_{CC}, must be reversed relative to the collector. Upon initial conduction through the E-B junction, holes are injected from the positive side of the 1.5-volt supply through the E-B junction and into the collector. Electrons from the 3-volt battery leave the negative side of the battery, flow through the lamp, and enter the top of the collector to recombine with holes.

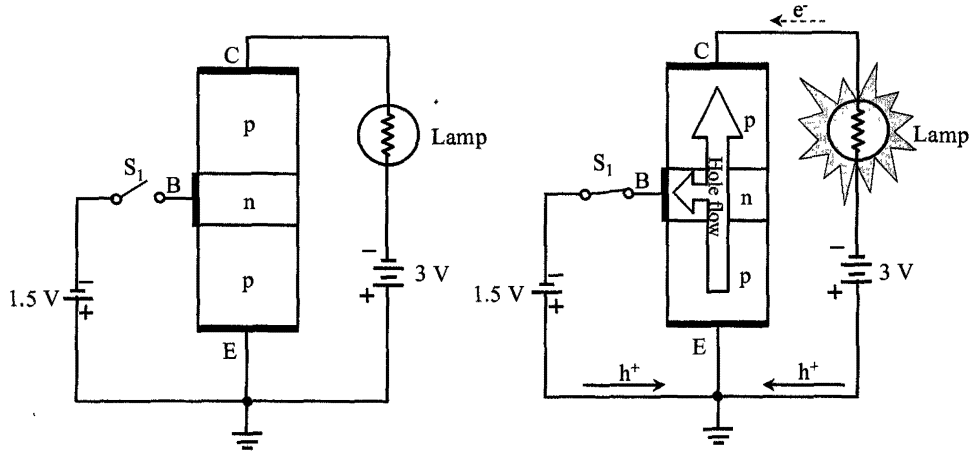


FIGURE 3.12 pnp Transistor Biasing Circuit

Structure of a Bipolar Junction Transistor ■ Figure 3.13 illustrates the structural features of the cross section of an npn bipolar junction transistor. Note the differences in the amount of dopant concentration in each of the electrodes. The emitter (E) and collector (C) are both heavily doped with n-type dopant, such as arsenic or phosphorus. The base (B) is lightly doped with boron, a p-type dopant. With fewer current carriers in the base, the current drawn through the base is significantly smaller than the current drawn through the collector. This difference accounts for the *gain* (amplification) in current from the input to the output of the transistor. This feature is the major difference between the transistor and the diode. The transistor is able to amplify small input signals literally hundreds of times to drive output devices such as speakers, motors, lamps, relays, and other electromechanical devices. The BJT is a *current-driven current amplifier* device. Thus, bipolar transistors are commonly used in radios, tape recorders, automotive electronics, aircraft control systems, biomedical instrumentation, robotics, and wherever there is a need for high power control.

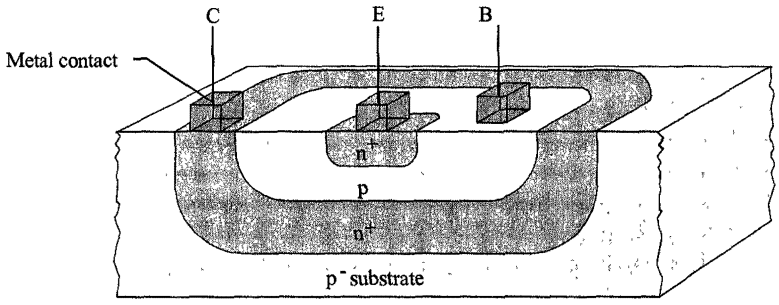


FIGURE 3.13 Cross Section of an npn BJT

Schottky Diode

The *Schottky diode* is formed when metal is brought in contact with lightly doped n-type semiconductor material. The resulting device operates just like an ordinary pn junction diode—low resistance when forward-biased and high resistance when reverse-biased. The forward junction voltage drop of the silicon Schottky diode (0.3 to 0.5 volts) is nearly half that of the silicon pn junction diode (0.6 to 0.8 volts). The Schottky diode's biggest advantage is that its conduction is due entirely to electrons, which results in faster switching time from on to off.[4] Figure 3.14 shows the schematic symbol and structured cross section of the Schottky diode.

The discovery of the Schottky diode has helped extend the usefulness of bipolar IC technologies into the twenty-first century. The Schottky diode concept has been utilized in the development of faster and more power efficient bipolar integrated circuits.

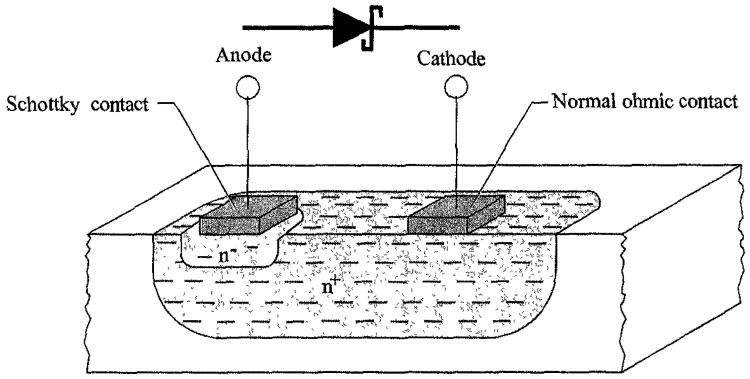


FIGURE 3.14 Schematic Symbol and Structural Cross Section of the Schottky Diode

Bipolar IC Technology

Diodes and bipolar transistors—along with supporting components of resistors, capacitors, insulators, and conductors—are used in the development of a breed of integrated circuits called *bipolar technology.* Bipolar technology was the first of many IC technologies that was used in the production of analog and digital integrated circuits. For many years bipolar devices have been noted for fast speeds, durability, and power-controlling abilities. The biggest drawback, however, has been their high power consumption. This, of course, meant higher costs for operating these devices in terms of electrical utilities and batteries for portable equipment. Some families of the bipolar technology era have since become obsolete. Other bipolar manufacturers have managed to maintain their business based on a continued need for bipolar devices for use in high-power applications. A list of the bipolar logic families is shown in Table 3.1. Further information about bipolar technology can be obtained by consulting the reference list at the end of this chapter.

TABLE 3.1 Bipolar Logic Families

Bipolar Logic Family	Abbreviation
Direct-Coupled Transistor Logic	DCTL**
Resistor-Transistor Logic	RTL**
Resistor-Capacitor-Transistor Logic	RCTL**
Diode-Transistor Logic	DTL**
Transistor-Transistor Logic*	TTL**
Schottky TTL Logic*	STTL†
Emitter-Coupled Logic*	ECL†

*Some forms of TTL, STTL, and ECL still in use through were 2000. **From G. Deboo and C. Burrous, *Integrated Circuits and Semiconductor Devices: Theory and Application,* 2nd ed. (New York: McGraw-Hill, 1977), p. 192. †From A. Sedra and K. Smith, *Microelectric Circuits* (Oxford: Oxford University Press, 1998), p. 1187, 1196.

CMOS IC Technology

When energy conservation became an international concern in the late 1970s, the semiconductor industry responded with the *field-effect transistor* (*FET*), which benefits from more compact and power-efficient electronic devices. Although early experiments with FETs date back to the 1930s, the first mass-produced FETs became available in the 1960s. Several versions of the FET have been used since then. Today the most popular IC technology, *CMOS* (*complementary metal-oxide semiconductor*), revolves around the improvements that have been made in FET design and manufacturing. The remainder of this text concentrates on the development and manufacturing processes related to CMOS devices.

The Field-Effect Transistor ■ The development of the field-effect transistor essentially created a new era in the history of the semiconductor industry. As opposed to the current-amplifying bipolar junction transistor, the FET is a *voltage-amplifying device.* The only similarities between the BJT and the FET are in the types of materials used to build the transistors and the number of electrodes. Both have three electrodes and both types are constructed of a single crystalline substrate. The greatest advantage of the FET is its low voltage and low power requirements. Whereas the BJT requires input current at the base to turn on the transistor, the FET turns on as a result of an electric field created when an input voltage is applied to the *gate*—thus the name field-effect transistor.

The FET has seen applications as an amplifier in linear/analog circuits and as a switching component in digital electronics. Its high input resistance and moderate gain characteristics make it an excellent device for use in instrumentation and communications. Its low power consumption and compactness make it extremely suitable for the ever-shrinking dimensions of VLSI and ULSI technology.

There are two basic types of FETs: the junction (*JFET*) and the metal-oxide (*MOSFET*) semiconductor. The major difference between these two types is that the gate on the MOSFET, which is the input to the FET, is insulated by a thin dielectric (silicon dioxide, referred to as *gate oxide*) from the other two electrodes of the transistor. The gate of the JFET actually forms a physical pn junction with the other electrodes of the transistor. JFETs are used extensively in GaAs integrated circuits. When metal gates are used in GaAs JFETs, the term *MESFET* is used. Because of the

popularity of MOSFETs in silicon VLSI applications, the remainder of our discussion of FETs will focus on the MOSFET. Consult the reference list at the end of this chapter for further information on the JFET.

MOSFETs. MOSFETs gained wide acceptance in logic applications in the 1970s and have been the mainstay transistor in IC products ever since. There are two categories of MOSFETs: *nMOS* (*n-channel*) and *pMOS* (*p-channel*). Each type is distinguishable by the majority current carriers in each device. Figure 3.15 shows some common schematic symbols and physical cross sections of the two types of MOSFETs.

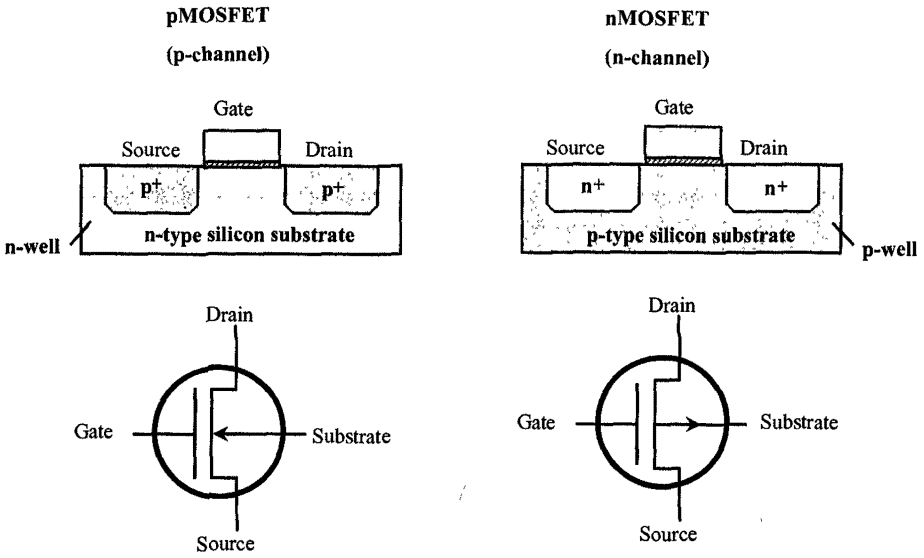


FIGURE 3.15 Two Types of MOSFETs

Each MOSFET has an input electrode called the gate. The term *metal oxide* refers to the material the gate is made of. However, the "metal" descriptor in MOS derives from the early days of MOS technology and is no longer true. The most popular material used in the formation of gates for MOSFETs is *polysilicon,* a polycrystalline silicon material that is deposited on the substrate during IC fabrication (see Chapter 11). The polysilicon, however, must be doped with one of the common p- or n-type dopants to give the material its conductive characteristics.

The *source* and the *drain* electrodes are heavily doped, respectively, with n-type or p-type dopant depending on the category of transistor being manufactured. The supply of majority current carriers comes from these two electrodes. The *nMOSFET* uses electrons as the majority carriers; therefore, the channel is n-type. The *pMOSFET,* then, has a p-channel formed by holes from the source and drain. When in the nonconduction mode, the channel is an open circuit consisting of an opposing doped region called the *well.* The *n-channel MOSFET* is built inside a *p-well,* while the *p-channel MOSFET* is built inside an *n-well.* When in the conduction mode, opposing carriers in the upper part of the well move away from the gate oxide interface and a channel of majority current carriers forms across from the source to the drain to complete the circuit (very similar to closing a switch).

Biasing the nMOSFET for Conduction Mode. Figure 3.16 on page 54 illustrates the biasing scheme for operating an n-channel MOSFET. The lamp in the circuit is an indicator of output current and normal operation of the transistor. At the moment there is no conduction in the n-channel MOSFET in Figure 3.16 because there is no input voltage applied to the gate. The condition of the channel is that of an open circuit. The majority of carriers in the area immediately beneath the gate oxide and between the source and drain electrodes are holes. At the moment the source and the p-well are at the same potential. As a pn junction there is no forward bias applied, so the pn junction remains in a nonconducting mode.

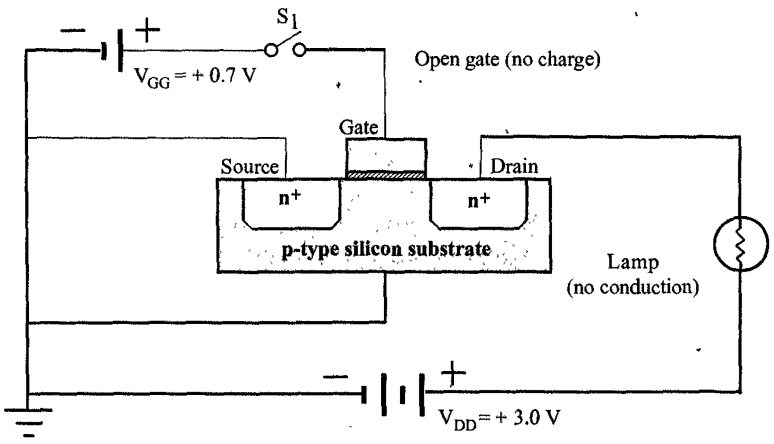


FIGURE 3.16 Biasing Circuit for an nMOS Transistor

When the switch, S_1, is closed as in Figure 3.17, certain conditions begin to happen. The small bias voltage of 0.7 V places a charge equal to the bias supply directly across the gate and source electrodes. An electrostatic charge is created with positive holes on the gate and electrons flowing up from the source to the gate oxide interface. At the same time an electric field is created by the positive charges at the gate. This field opposes the electric field of the holes in the upper part of the p-well. These opposing fields force the holes in the upper part of the p-well to move away from the gate oxide interface. As a result, electrons move from the source and drain to fill the gaps left vacant by the holes. The electric field from the positive charge on the gate continues to attract and hold the electrons in the channel area, thus closing the gap between the source and the drain with only electrons. Now, the electrons form a continuous crystalline structure of n-type silicon. Electrons from the 3-volt battery flow freely from the negative terminal through the source, to the n-channel, into the drain, and then through the lamp and back to the positive side of the battery. This condition will remain the same until any part of the input or output circuit is changed.

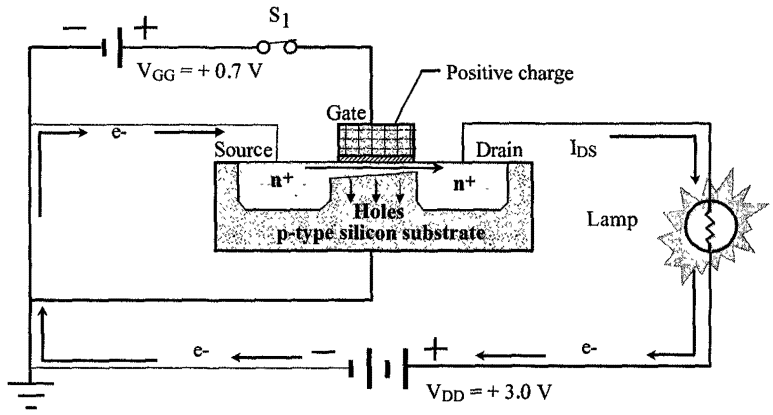


FIGURE 3.17 nMOS Transistor in Conduction Mode. Closing S_1 allows V_{GG} to place a positive charge on the gate of the transistor that causes an n-channel to form; therefore, allowing electrons to flow from the source to the drain.

Increasing the voltage at the gate increases all the electrical activities just mentioned. The electrostatic charge on the gate increases, which increases the electric field strength and forces the holes in the p-well to move further away from the gate-oxide interface. This results in an increase in the size of the n-channel and the number of electrons flowing through the channel. The net effect is for more drain current to flow, thus increasing the power delivered to the load (increased brilliance of the indicator lamp).

Figure 3.18 provides an example of a graphical representation of the electrical characteristics of an n-channel MOSFET. The graph represents drain current, I_{DS}, as a function of V_{DS} and V_{GS}. Each curve labeled V_{GS} represents a specific setting of the gate-source voltage. Using the curves it is possible to determine the value of I_{DS} for a given value of V_{GS} and V_{DS}. For example, assume V_{GS} = 4 V and V_{DS} = 2 V. The resulting drain current, I_{DS} = 0.28 mA. The *threshold voltage* is the lowest attainable V_{GS} value that will turn on the FET. This value is usually less than 1.0 V and is dependent on several variables, including structural sizes and doping characteristics of the electrodes and p-well.

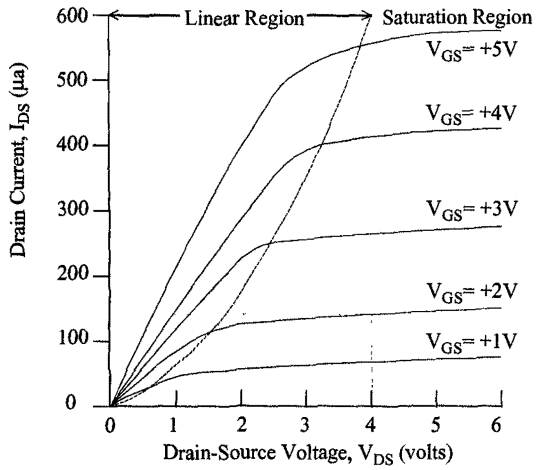


FIGURE 3.18 Example of Characteristics Curves of an n-channel MOSFET

Biasing the pMOSFET for Conduction Mode. Figure 3.19 illustrates the biasing circuit for a p-channel MOSFET. The operation of the p-channel transistor is similar to the operation of the n-channel type. The only difference is that the majority current carriers in the p-channel MOSFET are holes and the bias supplies are reversed. The other major difference between these two transistors is in terms of performance. The speed of the p-channel transistor is slower than that of its counterpart, primarily due to the holes, which move slower than electrons.

Closing S_1, allows V_{GG} to place a negative charge on the gate of the transistor that causes a p-channel to form, therefore allowing holes to flow from the source to the drain (see Figure 3.20 on page 56).

CMOS Technology ■ Manufacturers of MOSFET-based integrated circuits for many years concentrated on the development and manufacturing of products based solely on n-channel MOSFET technology. Whereas some semiconductor manufacturers produced both discrete nMOS and pMOS transistors, few, if any, manufactured integrated circuits made primarily with pMOS transistors.

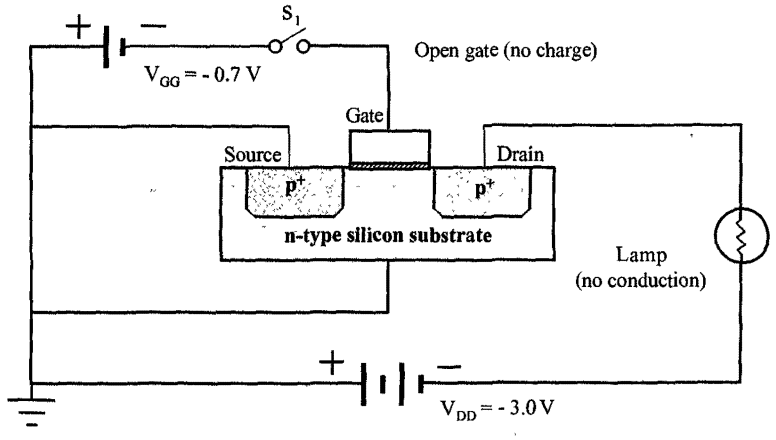


FIGURE 3.19 Biasing Circuit for a p-Channel MOSFET

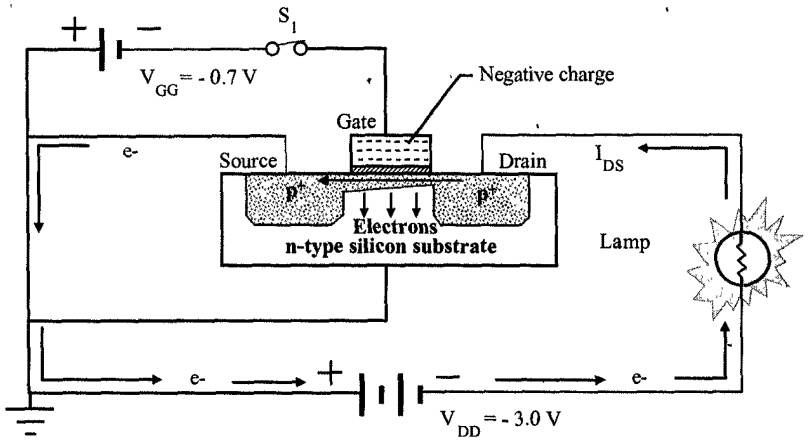


FIGURE 3.20 pMOS Transistor in Conduction Mode

Discrete pMOS transistors have served many usable functions in specific electronic applications, but in general, nMOS IC devices outperformed the pMOS technology. Thus, nMOS became the choice of most IC manufacturers.

CMOS incorporates both nMOS and pMOS transistors in the same integrated circuit. The combination of power efficiency, design scaling techniques, and improved manufacturing processes have made CMOS one of the most popular device technologies since the early 1980s. The term *scaling* has been commonly used to describe the process of *shrinking* the overall dimensions and operating voltages of existing ICs to attain improved operating performance as well as compactness. All dimensions and voltages must shrink in unison (thus the source of the word scaling) through the use of a design model employed by the IC designer during the circuit design and layout stage.

A schematic diagram of a simple CMOS inverter circuit is shown in Figure 3.21. The schematic shows the gates of both transistors connected together. These gates serve as a single input to the inverter. The output of the inverter is taken out of the two drains, which are tied together. The source of the n-channel transistor is grounded, while the source of the p-channel transistor is connected to the bias supply, V_{DD}. A signal applied to the input of the CMOS is inverted at the output, as shown in Figure 3.21. Normal operation has the n-channel operating during the positive transition of the input signal and the p-channel operating during the negative transition. The efficiency of the CMOS inverter circuit occurs during the transition period when the input signal is at zero. There is no power consumed by the transistors when the input signal is at zero. nMOS, TTL, and ECL circuits differ from CMOS in that those logic families all dissipate power even with no signal applied. This fact is one major reason why CMOS is currently the preferred IC technology for use in the manufacture of portable electronics products such as calculators, clocks, cellular telephones, and notebook computers.

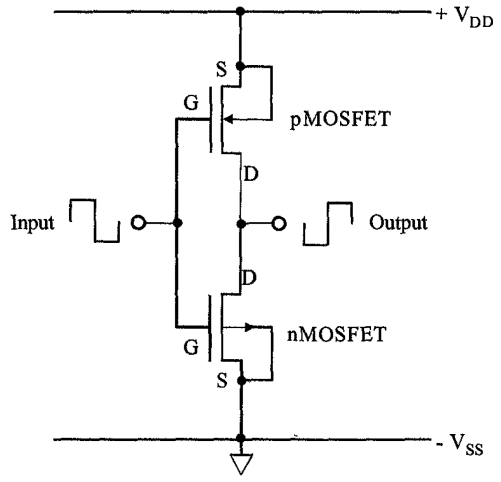


FIGURE 3.21 Schematic of a CMOS Inverter

The physical structure of a simple CMOS inverter is shown in the top view in Figure 3.22 and cross-sectional view in Figure 3.23. The n-substrate serves as the n-well for the p-channel transistor. A separate p-well must be created for the n-channel transistor. This well is created by a selective doping process called *ion implantation*. Isolation regions separate transistors from each other and from other transistors (not shown). The isolation regions are referred to as *field oxide,* which is manufactured from silicon dioxide (also called *glass*). The field oxide insulates transistors from each other in order to prevent the flow of undesirable leakage current between transistors. The actual electrical connections between the two transistors are made possible with metal deposited over the insulated transistors. Later, a patterning process is used to define the exact locations where the metal connections are to be made, followed by an etching process to remove the excess metal. Further details of the CMOS manufacturing process are beyond the scope of this chapter. Chapter 9 focuses on the specific physical and chemical requirements and process flow in CMOS IC manufacturing.

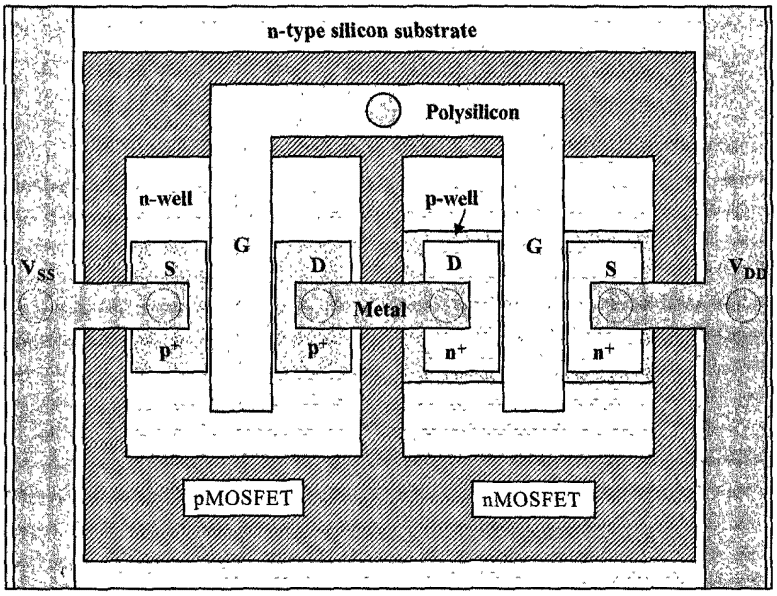


FIGURE 3.22 Top View of CMOS Inverter

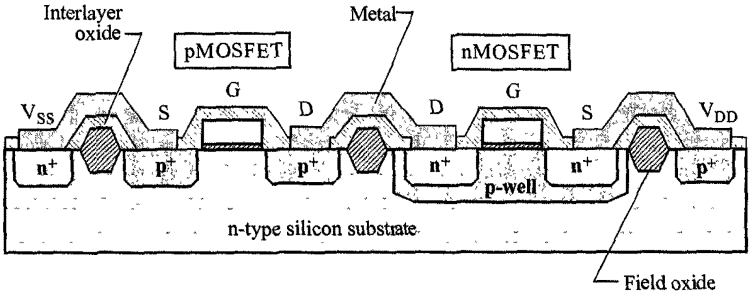


FIGURE 3.23 Cross-section of CMOS Inverter

BiCMOS Technology ■ *BiCMOS technology* makes use of the best features of both CMOS and bipolar technology in the same IC device. BiCMOS incorporates the low-power, high-density CMOS structures with the high-current drive capabilities of TTL or ECL device structures. Applications of BiCMOS products can be found wherever the need for complex digital control of high-power loads is desired. In which case, *digital/analog* (D/A) converter chips may be used to provide the analog drive signals that are used to control electromechanical equipment. On the instrumentation side, *analog/digital* (A/D) chips may be used to measure the outcome of analog drive

signal. Figure 3.24 shows a basic example of BiCMOS chips used in an instrumentation and control application. Other applications for BiCMOS include automotive electronics, aerospace, robotics, and industrial equipment.

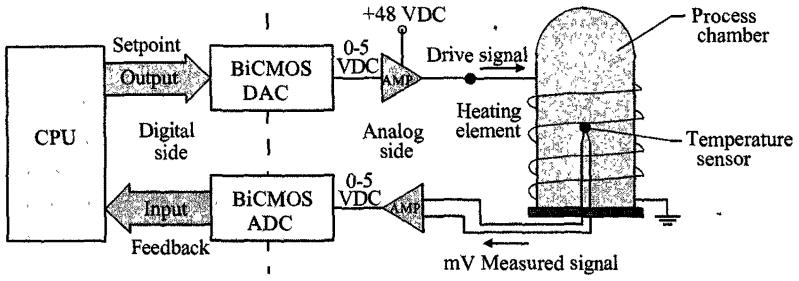


FIGURE 3.24 BiCMOS Chips used in the Control of a Simple Heating System

Figure 3.25 shows additional bipolar transistors added to the CMOS inverter mentioned earlier in Figure 3.21 on page 56. The bipolar transistors (Q3, Q4) provide greater current drive than the CMOS inverter (Q1, Q2) is capable of.

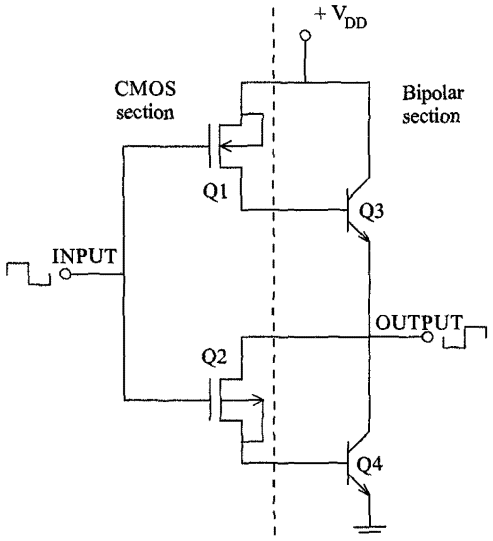


FIGURE 3.25 Simple BiCMOS Inverter
Redrawn from H. Lin, J. Ho, R. Iyer, and K. Kwong, "Complementary MOS-Bipolar Transistor Structure," *IEEE Transactions Electron Devices,* ED-16, 11 Nov. 1969, p. 945–951.

Enhancement and Depletion-Mode MOSFETs

Another way to categorize FETs is in terms of *enhancement mode* and *depletion mode.* Up to now all n-channel and p-channel MOSFETs discussed in this chapter have been enhancement-mode transistors. Enhancement-mode MOSFETs are the most commonly used transistors in the industry. The difference between the enhancement-mode and depletion-mode transistors is shown in Figure 3.26.

In general, the major difference between enhancement- and depletion-mode transistors is in the way the channels are doped. While the enhancement-mode channels are doped opposite in polarity to the source and drain regions, the depletion-mode channels are doped the same as their respective source and drain regions. This setup creates open and closed circuit conditions, respectively, for the enhancement- and depletion-mode transistors. Consequently, the enhancement-mode transistor is regarded as *normally-off* and the depletion mode as *normally-on.*

The enhancement-mode transistor works very well in digital logic applications and requires only a single polarity input signal (V_{GS}) to operate the FET. On the other hand, the depletion-mode transistor is partly turned on by the already existing closed channel. The input voltage can swing in one direction to increase the current through the channel or reverse in the opposite direction to decrease the current through the channel. The depletion-mode transistor will turn off completely if the input voltage to the gate is further increased in the opposite direction. Both types of transistors can be designed into practical applications in both analog and digital circuit applications in either discrete component applications or in integrated circuit applications.

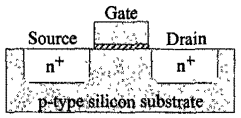
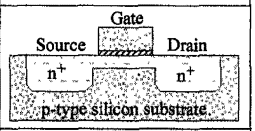
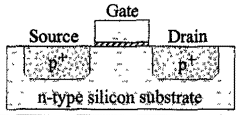
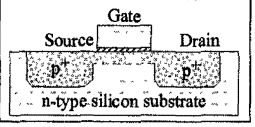
MOSFET Type	Mode	Standby Condition	V_{GG} Switching Requirements	Physical Structure
nMOS	Enhancement	Off	+	
nMOS	Depletion	On	-	
pMOS	Enhancement	Off	-	
pMOS	Depletion	On	+	

FIGURE 3.26 Comparison of Enhancement- and Depletion-Mode MOSFETS

LATCHUP IN CMOS DEVICES

Just as undesirable parasitic resistances and capacitances exist in semiconductor devices, sometimes the pn junctions in CMOS devices can produce *parasitic transistors* that can create a *latchup* condition in CMOS ICs that causes transistors to unintentionally turn on. Figure 3.27 illustrates the parasitic transistors in a CMOS inverter structure. Complementary junction transistors are formed as a result of normal manufacturing of MOSFETs in a CMOS structure. Given certain operating conditions it is possible for the parasitic junction transistors to turn on and create a low resistance path for current to flow across the CMOS structure.[5] The transistors become latched, thus preventing any further control of the MOSFETs in the CMOS device.

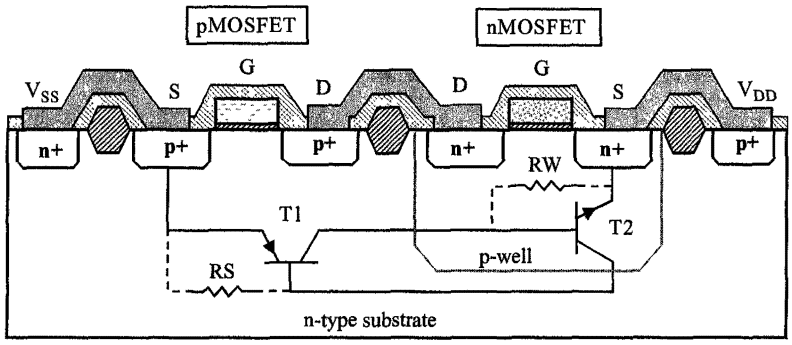


FIGURE 3.27 Parasitic Junction Transistors within a CMOS Structure

Indeed, the latchup phenomenon is a very complex concept. Any further discussion is beyond the scope of this chapter. Suffice it to say that knowing about its existence will help you understand the design and manufacturing steps that are taken to prevent this condition. Later chapters describe several manufacturing techniques used in the prevention of latchup, namely creating isolation barriers between transistors, placing an epitaxial layer explained in Chapter 11 between the substrate and the CMOS structures, and using ion implantation to create retrograde wells explained in Chapter 17.

INTEGRATED CIRCUIT PRODUCTS

Semiconductor products can be found in most electrical and electronic apparatuses throughout the world. The most familiar applications of ICs are found in the personal computer (PC). Such a computer would employ a microprocessor unit (MPU), read-only memory (ROM), random-access memory (RAM), communications interface chips, video graphics controller chips, disk drive controller chips, and more. For the most part, these chips are built using CMOS technology, followed by BiCMOS and bipolar technologies.

The types of ICs used vary depending on customer needs, environment, cost, power requirements, and other specifications. Most ICs, as was mentioned earlier, fall into two basic categories —analog or digital. The following sections list the most common types of ICs found in most consumer and industrial applications.

Linear IC Product Types

The *linear IC* family consists of devices that are designed to operate primarily in analog circuit applications. Linear ICs are typically used in audio systems, radio communications, industrial controls and instrumentation, aerospace, and in automotive electronics. The following are a few examples of linear ICs.

Operational Amplifier ■ The *op-amp* is a high-gain and high-input impedance amplifier that can be adapted to a variety of electronic control applications. It can be used as an amplifier, oscillator, voltage regulator, voltage limiter, sample-and-hold amp, voltage rectifier, integrator, and differentiator.

Voltage Regulator ■ The *voltage regulator* is a single-chip device containing diodes, transistors, resistors (VR) and capacitors for the sole purpose of regulating the voltage delivered to a load. VRs can be used in any electronic subassembly or system where a constant voltage source is needed to maintain a continuous voltage over a specific range of load resistances. Some applications include computers, peripherals, and various types of instrumentation.

Stepper Motor Driver ■ The *stepper motor driver* (*SMD*) is a bipolar IC used for controlling stepper motors. An SMD can be used any place where a stepper motor must be accurately controlled, such as in laser printers, photocopiers, scanners, and robotics, as well as aerospace, automotive, and industrial applications.

Digital IC Product Types

The digital IC family includes devices that operate with binary (1s and 0s) bits of data signals, such as those used in calculators, computers, digital pagers, cellular phones, and many others similar applications.

Volatile Memory ■ *Volatile memory* is a semiconductor device that allows data to be stored and changed as desired. Data in the volatile memory is lost when the power is turned off. Volatile type memory devices are used in computers, calculators, and appliances as well as automotive, aerospace, medical, military, and industrial equipment—in any application where logic instructions must be stored and changed as needed by the user. All memory devices contain millions of individual memory cells. The probability of defects increases with device density; therefore, redundant

memory cells are built into memory chips. Semiconductor fuses are built into the memory cells that are deliberately blown during the initial product test to disable defective cells.

RAM. A *random-access memory* device is one that can be accessed to read its stored data or it can be erased and new data rewritten into it. RAM chips do not need to be removed from the printed circuit board to change the data held in storage within its memory cells. Its contents can be easily changed in the normal operations of the logic system.

DRAM. The *dynamic RAM* is the most common and least expensive of the RAM family. The term *dynamic* refers to the refresh voltage that must be applied regularly to the storage capacitors to retain the data. DRAMs require more power to operate the capacitors.

SRAM. The *static RAM* uses flip-flops as data storage registers. SRAMs do not require refreshing; thus, they use less power than the DRAMs. The data is also lost when the power is removed.

MPU or CPU. *Microprocessor units* (also known as *central processing units*), are complex logic ICs that are capable of executing instructions programmed into a separate or internal ROM. The MPU is capable of decision making and of performing math functions. Microprocessors are used in computers, calculators, and appliances as well as automotive, aerospace, medical, military, and industrial equipment—in any application where a controlling function is needed or where computer functions are a requirement.

Nonvolatile Memory ■ The *nonvolatile memory* is a semiconductor device designed to store digital data in the form of an electrical charge. The charge remains in storage even after the power is turned off. Nonvolatile type memory devices are used in computers, calculators, and appliances as well as automotive, aerospace, medical, military, and industrial equipment—in any application where logic instructions must be stored for later recall.

ROM. A *read-only memory* IC is nonvolatile and is programmed directly in the IC manufacturing process. It is also referred to as a *mask-programmable* ROM. The mask sets used during manufacturing contain code patterns for a specific ROM. ROMs are very expensive to produce. ROMs were the earliest version of nonvolatile memory devices. The consumer can read electronically what has been coded into the ROM but cannot change its contents.

PROM. A *programmable read-only memory* is an IC that can be reprogrammed in the field and is less expensive than the mask-programmable ROM. This tool applies large voltage pulses that allow memory cells to be changed as needed. The data programmed into this memory cannot be changed during normal operation of a computer. There are two types of *field-programmable* proms —the EPROM and the EEPROM.[6]

EPROM. An *erasable PROM* can be erased and reprogrammed in the field. The chip must be removed from the circuit board and erased by exposure to UV light. The EPROM can then be reprogrammed. The EPROM is an improvement over its two predecessors; however, any change in the stored code requires complete erasure of the PROM followed by reprogramming.

EEPROM. The *electrically erasable PROM* IC can be electrically erased and reprogrammed in the field without having to remove it from the circuit board. The EEPROM is the most convenient form of ROM. The device operates more like a random-access memory in that it can be erased and rewritten without the use of any special equipment. EPROMs have many useful applications in the computer and peripheral industry.

Flash RAM. Flash memory is a type of nonvolatile memory that can be erased and reprogrammed—similar to electrically-erasable programmable read-only memory (EEPROM), but faster to update. Flash memory is often used to store the operating code such as the Basic Input/Output System (BIOS) in a personal computer. Flash memory is used in computers, digital cellular phones, digital cameras, embedded controllers, and others products.

ASIC. *Application-specific ICs* are fully custom-designed and manufactured to meet an individual customer's needs. The ASIC device may include existing logic circuitry with new design features as requested by the customer. ASIC chips are also nonvolatile. All logic functions of the IC

are determined by specific masks that have been designed to achieve unique structural features given in the specific ASIC design specifications. While ASIC devices provide the exact product to meet the user's needs, the cost of manufacturing them is high.

PLD. *Programmable logic devices* are ICs that utilize a wide variety of logic components. The actual logic function implemented is determined by the user, using some form of design software to specify the state of the internal programming points.[7] During the programming process, higher than normal voltages are applied to specific semiconductor fuses. The fuses heat up and vaporize leaving behind a desirable logic circuit configuration as programmed by the user.

PAL. A *programmable array logic* IC contains a network of programmable logic gates used to create a custom logic circuit. The PAL has an array of input AND gates that drive an array of hard-wired output OR gates. The AND gates can be field-programmed as required, but the output OR gates are fixed.[8]

PLA. A *programmable logic array* differs from the PAL in that both the input AND gates and the output OR gates are programmable. PLAs are available in both mask-programmed and field-programmed versions.

MPGA. A *mask-programmable gate array* chip can be customized to meet the functional needs of individual customers. Initially during the manufacturing process all gate array products passing through a specific manufacturing facility may contain the same basic number and configuration of transistor circuitry. It is during the interconnect layers of the manufacturing process that specific masking steps are used to define the function of thousands of transistors within individual chips. The mask sets are designed in cooperation with customers to define the proper interconnects for a specific gate-array application.[9]

FPGA. A *field-programmable gate array* is an IC device in which the final logic product can be defined in the field without the use of an IC manufacturing facility.[10] Programmable logic switches within the FPGA can be activated to create connections between desired logic circuits. The convenience of customizing in the field and its low manufacturing cost make the FPGA a viable alternative to the MPGA. However, the MPGA device operates faster than the FPGA since all interconnects are hard-wired internally during the IC manufacturing process.

SUMMARY

Electronic components are assembled into analog or digital circuits. Passive components, such as resistors and capacitors, conduct current without specific polarity requirements and are constructed on silicon substrates. These components have parasitic losses that are detrimental to IC performance. Active components, such as transistors and diodes, control the direction of current flow. The pn junction diode can form an open circuit (reversed biased) or closed circuit (forward biased). The bipolar junction transistor (BJT) can be biased to conduct current and also amplify input signals. The structure of the BJT is based on n-type or p-type dopant regions that form two pn junctions per transistor as either an npn or pnp. The Schottky diode has faster switching times, which has helped bipolar IC technology extend its applications in to the twenty-first century.

CMOS IC technology is more power efficient than bipolar and is one of the most popular device technologies with the most common transistor being the MOSFET as either an nMOS or pMOS. The MOSFET has three electrodes, the gate, source and drain that are attached to a biasing circuit for conduction mode. The minimum voltage to turn on a MOSFET is the threshold voltage. An example of CMOS technology is the inverter circuit. BiCMOS technology combines both CMOS and bipolar technology. MOSFETs can be in enhancement (normally-off) or depletion (normally-on) mode for both analog and digital applications. Parasitic transistor effects can occur in pn junctions that create a latchup condition that causes the transistor to turn on with no further control. There is a wide range of IC products, including analog and digital.

KEY TERMS

structure
analog circuit
linear circuit
digital circuit
logic devices
device technology
passive components
parasitic resistance
planar capacitors
parasitic capacitance
active components
pn junction diode
carrier-depletion region
barrier voltage
reverse bias
forward bias
bipolar junction transistor (BJT)
npn
pnp
emitter
base
collector
current-driven current amplifier
well
n-channel MOSFET
p-well
p-channel MOSFET
n-well
scaling
shrinking
field oxide
glass
BiCMOS technology
digital/analog (D/A)
analog/digital (A/D)
enhancement mode
depletion mode
normally-off
normally-on
parasitic transistors
latchup
op amp
voltage regulator (VR)
stepper motor driver (SMD)
volatile memory
random-access memory (RAM)
gain
Schottky diode
bipolar technology
field-effect transistor (FET)
CMOS (complementary metal-oxide semiconductor)
voltage-amplifying device
gate
JFET
MOSFET gate oxide
MESFET
nMOS (n-channel)
pMOS (p-channel)
polysilicon
source electrodes
drain electrodes
nMOSFET
pMOSFET
dynamic RAM (DRAM)
static RAM (SRAM)
microprocessor units (MPU)
central processing units (CPU)
nonvolatile memory
read-only memory (ROM)
programmable read-only memory (PROM)
field-programmable PROM
erasable PROM
electrically erasable PROM
flash RAM
application specific IC
programmable logic devices (PLD)
programmable array logic (PAL)
programmable logic array (PLA)
mask-programmable gate array (MPGA)
field-programmable gate array (FPGA)

REVIEW QUESTIONS

1. What is an analog circuit? Provide two examples of applications.
2. What is a digital circuit? Provide two examples of applications.
3. What is a passive component? Give two examples of this type of component.
4. List three ways to make a resistor on a single crystalline substrate.
5. What is parasitic resistance? Why is it undesirable in ICs?
6. What is a planar capacitor? Describe four techniques for building this component on a silicon substrate.
7. What is parasitic capacitance? What problems can this condition create in ICs?
8. What is an active component? Give two examples of this type of component.
9. When is a pn junction diode formed? Why is this junction important for an IC?
10. What is the barrier voltage of a pn junction, and how is it formed?
11. Explain what happens when a pn junction is reversed biased.
12. Explain what happens when a pn junction is forward biased.
13. A bipolar junction transistor (BJT) has how many electrodes, junctions, and variations? What are the names of the electrodes? What are the names of the variations?
14. Refer to Figure 3.11. What happens when switch S1 is open? What happens to the E-B junction when switch S1 is closed? When does conduction occur and why?
15. Describe the dopant concentrations in the emitter, collector, and base in an npn bipolar junction transistor.
16. What type of amplifier device is the BJT? How does this affect the applications in terms of power requirements?
17. How is the Schottky diode formed? What is its biggest advantage?
18. What are some notable characteristics of bipolar technology? What is biggest drawback to bipolar technology?
19. What are the benefits of the field-effect transistor (FET)?
20. What is the most popular IC technology today?
21. Is the FET a voltage-amplifying or current-amplifying device?
22. What is the greatest advantage of the FET?
23. How does the FET turn on? What voltage level must be applied to turn on the FET and where should the voltage be applied in the FET?
24. Why is the FET suitable for VLSI and ULSI technology?
25. What are the two basic types of FETs? What is the major difference between them?
26. What are the two categories of MOSFETs? How are they distinguishable from one another?
27. What is the most popular conductor material used in the formation of gates for MOSFETs? How is this material made a conductor?
28. What purpose does the source and drain serve in a MOSFET?
29. What is the majority carrier for a nMOSFET, and what type of channel does it have?
30. What is the majority carrier for a pMOSFET, and what type of channel does it have?
31. What is the well in a MOSFET? Explain what happens in the well region in the conduction mode.
32. If an nMOSFET is in an open-circuit mode, explain what is happening in the three electrode regions (i.e., the source, the drain and the channel below the gate).
33. When switch S1 is closed on the nMOSFET transistor as shown in Figure 3.17, explain what happens with respect to holes and electrons in the electrode regions.
34. What is the threshold voltage in a FET?
35. What is the major difference in performance between a pMOSFET and nMOSFET?
36. What does scaling mean with regards to IC design?
37. What is the efficiency of a CMOS inverter circuit with respect to power consumption?
38. If a p-substrate is used for a CMOS inverter, how is the n-well created?
39. What is the purpose of the field oxide? What is another name for silicon dioxide?
40. What two IC technologies are used in BiCMOS?
41. What could a digital/analog (D/A) converter chip be used for? What could an analog/digital (A/D) chip be used for?
42. Explain the difference between an enhancement-mode transistor and depletion-mode transistor with regards to their standby condition.
43. Which type of MOSFET, enhancement-mode or depletion-mode, is the most commonly used in the semiconductor industry?
44. What is latchup in CMOS devices? What undesirable condition can cause this?
45. Describe three manufacturing techniques used to prevent latchup. What is a linear IC and when would it be used?
46. What is an op amp?
47. What is a voltage regulator?
48. What is a stepper motor driver?
49. What is volatile memory?
50. Describe RAM, DRAM, and SRAM memory.
51. What is MPU?
52. What is nonvolatile memory?
53. Describe the differences among ROM, PROM, EPROM, and EEPROM memory. What is flash RAM, and what are some applications?
54. What is ASIC?
55. What is PLD?
56. What are PAL and PLA?
57. What is an MPGA and a FPGA?

IC MANUFACTURERS' WEB SITES

Actel	http://www.actel.com/
Advanced Micro Devices	http://www.amd.com/
Altera	http://www.altera.com/
Analog Devices	http://www.analog.com/
AT&T Tech. History	http://www.akpublic.research.att.com/
Burr Brown	http://www.bbrown.com/
Cirrus Logic	http://www.cirrus.com/
Cypress Semiconductor	http://www.cypress.com/
Dallas Semiconductor	http://www.dalsemi.com/
Fairchild Semiconductor	http://www.fairchildsemi.com/
Fujita Laboratory	http://www.fujita3.iis.u-tokyo.ac.jp/
Fujitsu	http://www.fujitsu.com/
General Semiconductor	http://www.gensemi.com/
Hitachi Semiconductor	http://semiconductor.hitachi.com/
IBM Microelectronics	http://www.chips.ibm.com/
Intel	http://www.intel.com/
International SEMATECH	http://www.sematech.org/public/index.htm
Intersil Corporation	http://www.intersil.com/
Lattice Semiconductor	http://www.latticesemi.com/
LSI Logic	http://www.lsilogic.com/
Micron Semiconductor	http://www.micron.com/mti/
Mitsubishi Silicon Amer.	http://www.mmc-sil.com/
Motorola Semiconductor	http://mot-sps.com/
National Semiconductor	http://www.national.com/
NEC Semiconductor	http://www.nec.com/semiconductors/
Philips Semiconductor	http://www-us2.semiconductors.philips.com/
QuickLogic	http://www.quicklogic.com/
Rockwell International	http://www.rockwell.com/
Samsung Semiconductor	http://www.usa.samsungsemi.com/
SEMATECH	http://www.sematech.org/public/index.htm
ST Microelectronics	http://www.st.com/stonline/index.shtml
Texas Instruments	http://www.ti.com/
Vishay Siliconix	http://www.vishay.com/brands/siliconix/
Xicor	http://www.xicor.com/
Xilinx	http://www.xilinx.com/
Zilog	http://zilog.com/

Note: Web site addresses are subject to change. The Web site addresses listed here were current at the time of publication.

REFERENCES

1. S. Wolf, *Silicon Processing for the VLSI Era,* vol. 2 of *Process Integration,* IC Resistor Fabrication, (Sunset Beach, CA: Lattice Press, 1990), Appendix A.
2. R. Jacob Baker, H. Li, and D. Boyce, chap. 7 in *CMOS Circuit Design, Layout and Simulation,* (IEEE Press, 1998).
3. A. Sedra and K. Smith, *Microelectronic Circuits,* (New York: Oxford University Press, 1998), p. 138.
4. Ibid., p. 197.
5. R. Jacob Baker, H. Li, and D. Boyce, *Circuit Design,* p. 212–215.
6. S. Brown et al., *Field-Programmable Gate Arrays,* (Norwall, MA: Kluwer Academic Publishers, 1992), p. 2.

7. R. Seals and G. Whapshott, *Programmable Logic PLDs and FPGAs,* (New York: MacMillan Press Ltd., 1997), p. 1.
8. J. Carter, *Digital Designing With Programmable Logic Devices,* (Upper Saddle River, NJ: Prentice-Hall, 1997), p. 17.
9. M. Smith, *Application-Specific Integrated Circuits,* (Reading, MA: Addison-Wesley Longman, 1997), p. 4–15.
10. S. Brown et al., *Field Programmable Gate Arrays,* (Norwell, MA: Kluwer Academic Publishers, 1992), p. 1.

CHAPTER 4

SILICON AND WAFER PREPARATION

The primary semiconductor material used to fabricate microchips is silicon, the most important material in the semiconductor industry. In order for silicon to be acceptable for the fabrication of semiconductor devices, a special purity grade is used that meets stringent material and physical conditions.

The typical U.S. semiconductor manufacturer does not produce silicon wafers. The production of the silicon material and its preparation into wafers occurs in a highly specialized factory dedicated to the production of silicon wafers. These wafers are then supplied to wafer fabrication facilities for processing into various types of microchips by semiconductor manufacturers.

The quality of the final microchip produced in a wafer fab depends directly on the quality of the starting silicon wafer. If defects exist in the incoming wafer, then there will surely be defects in the microchip. An understanding of the silicon wafer and its preparation process will assist us in comprehending the silicon wafer's importance to the complete microchip process.

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Describe how raw silicon is refined into semiconductor-grade silicon.
2. Explain the crystal structure and growth method for producing monocrystalline silicon.
3. Discuss the major defects in silicon crystal.
4. Outline and describe the basic process steps for wafer preparation, starting from a silicon ingot and finishing with a wafer.
5. State and discuss seven quality measures for wafer suppliers.
6. Explain what epitaxy is and why it is important for wafers.

INTRODUCTION

To exploit silicon's desirable properties as a semiconductor, the natural silica must be purified into an extremely clean silicon material. Pure silicon is required to minimize microdefects at the atomic level of the silicon that are detrimental to semiconductor performance. Once pure silicon is obtained, it is then transformed into the physical wafer with the desired crystal orientation, proper amount of dopants, and physical dimensions necessary for semiconductor wafer fabrication.

SEMICONDUCTOR-GRADE SILICON

The highly refined silicon used for wafer fabrication is termed *semiconductor-grade silicon* (SGS), which is also sometimes referred to as electronic-grade silicon. Obtaining SGS from natural silicon with the acceptable purity level needed for semiconductor device performance is a multistep process (see Table 4.1).[1] There are alternative methods for obtaining SGS, but the following description is the predominant production process.

TABLE 4.1 Steps to Obtaining Semiconductor-Grade Silicon (SGS)

Step	Description of Process	Reaction
1	Produce metallurgical-grade silicon (MGS) by heating silica with carbon.	$SiC\ (s) + SiO_2\ (s) \rightarrow Si\ (l) + SiO(g) + CO\ (g)$
2	Purify MG silicon through a chemical reaction to produce a silicon-bearing gas of trichlorosilane ($SiHCl_3$).	$Si\ (s) + 3HCl\ (g) \rightarrow SiHCl_3\ (g) + H_2\ (g) + heat$
3	Using the Siemens process, $SiHCl_3$ and H_2 react to produce pure semiconductor-grade silicon (SGS).	$2SiHCl_3\ (g) + 2H_2\ (g) \rightarrow 2Si\ (s) + 6HCl\ (g)$

The first step to achieving SG silicon is to produce a metallurgical-grade silicon (MGS) by heating silica (SiO_2), a pure sand, with carbon in a reducing-gas atmosphere.

$$SiC\ (solid) + SiO_2\ (solid) \rightarrow Si\ (liquid) + SiO\ (gas) + CO\ (gas)$$

The resulting metallurgical-grade silicon on the right side of the reaction is 98% pure. Because MG silicon has unacceptable levels of contaminants, it is of no use for semiconductor fabrication. This metallurgical-grade silicon is then crushed and purified through a chemical reaction that produces a silicon-bearing gas of trichlorosilane ($SiHCl_3$).

$$Si\ (solid) + 3HCl\ (gas) \rightarrow SiHCl_3\ (gas) + H_2\ (gas) + heat$$

The silicon-bearing gas of $SiHCl_3$ undergoes additional chemical processing and reduction in hydrogen to produce 99.9999999% (nine nines) pure, semiconductor-grade silicon (SGS).[2] This rate equation is,

$$2SiHCl_3\ (gas) + 2H_2\ (gas) \rightarrow 2Si\ (solid) + 6HCl\ (gas)$$

The pure SGS is produced in a process referred to as the Siemens process.[3] The $SiHCl_3$ gas and hydrogen are injected into the Siemens reactor (See Figure 4.1) and chemically react on heated rods of ultrapure silicon (the rod temperature is about 1100°C). After several days, the process is complete and the rods of deposited SGS are cut up into smaller pieces for growth into the silicon crystal.

Semiconductor-grade silicon has the ultrahigh purity required for semiconductor manufacturing, with less than two parts per million (ppm) of carbon and less than one part per billion (ppb) for elements from Groups III and V (the critical doping elements) of the periodic table.[4] However, the silicon produced by the Siemens process does not have atoms arranged in a desirable crystal order. We will now analyze crystal structure to understand the correct atomic order for semiconductor-grade silicon.

CRYSTAL STRUCTURE

It is not only critical that the semiconductor-grade silicon used to fabricate semiconductor devices have ultrahigh purity, but it must also have a near-perfect crystal structure. This condition is necessary to avoid electrical and mechanical defects that are detrimental to device performance.

A *crystal* is a solid material that has an ordered, repeatable three-dimensional pattern over a long range of many atoms. Figure 4.2 shows how atoms are related in a crystal structure, referred to as the *crystal lattice.* The crystal lattice represents the repeatable order within a crystal at the atomic level of the internal structure (atomic order), even though the surfaces of the crystal may be abraded or rough. An example is the uneven surface of common beach sand, which has an internal crystalline structure with atomic order. Our goal is to obtain a specific crystal lattice necessary for the specific requirements for SG silicon used to fabricate wafers.

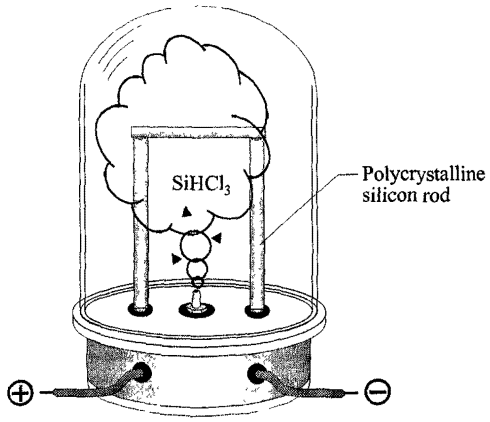


FIGURE 4.1 Siemens Reactor for SG Silicon

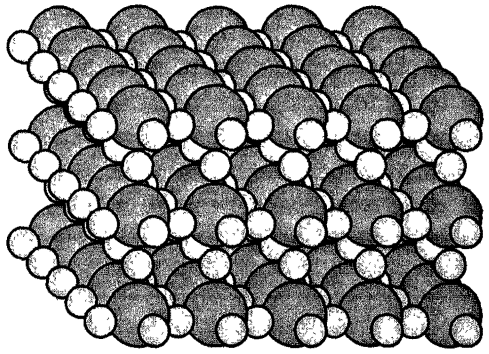


FIGURE 4.2 Atomic Order of a Crystal Structure

Amorphous Materials

Amorphous materials are noncrystalline solids that lack a repetitive structure and demonstrate structural disorder at the atomic level (see Figure 4.3). Amorphous materials are the opposite of crystalline materials. Plastic is an example of an amorphous material. Amorphous silicon would be an unacceptable structure for the silicon used in wafers to fabricate semiconductor devices. This is because many of the electrical and mechanical aspects of a device occur at the atomic level in the silicon, which requires order and predictability to yield repeatable results from chip to chip.

Unit Cells

The most fundamental entity for the long-range order found in the atomic pattern of a crystal material is the *unit cell.* A unit cell is the simplest arrangement of atoms that, when repeated in a three-dimensional framework, gives the crystal structure. We need to study unit cells in order to understand how silicon replicates itself in its crystal structure. A two-dimensional analogy to a unit cell can be seen in a checkerboard or a section of a tiled floor. A three-dimensional analogy would be a neat arrangement of a child's building blocks. An illustration of a unit cell in a three-dimensional structure is shown in Figure 4.4.

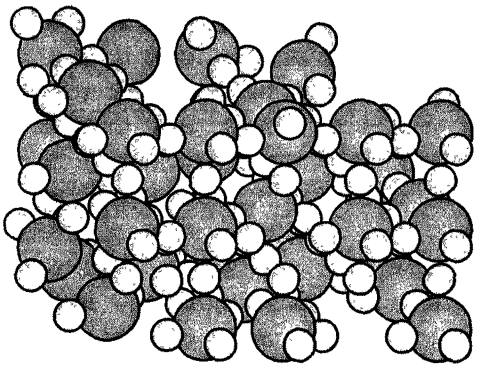


FIGURE 4.3 Amorphous Atomic Structure

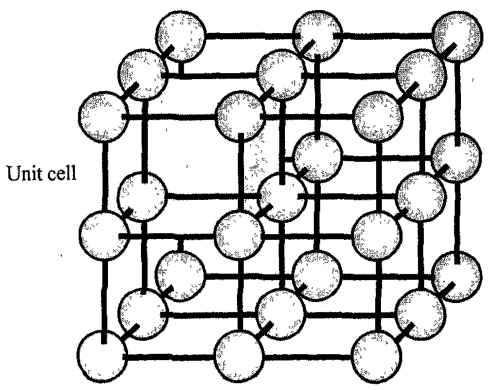


FIGURE 4.4 Unit Cell in 3-D Structure

Since crystal structures are identical in three dimensions, unit cells have a framework structure, such as a cube. There are seven possible crystal systems in nature.[5] The crystal system relevant to silicon technology is the cubic system. We will only consider one of the basic crystal structures because it is applicable to silicon: *face-centered cubic,* or *FCC.* The FCC unit cell is shown in Figure 4.5.

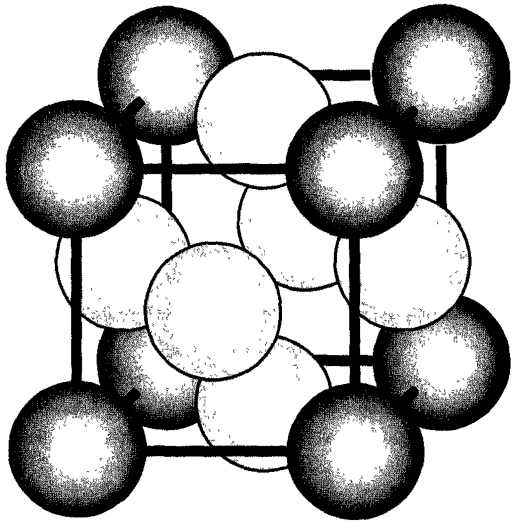


FIGURE 4.5 Face-Centered Cubic (FCC) Unit Cell

Within a crystal structure, unit cells are tightly packed and, therefore, share atoms. Sharing atoms is important because this is how unit cells build up into a cohesive crystal lattice structure. In an FCC unit cell, each corner shares an atom with eight other unit cells, while each face shares with one other unit cell. For an FCC unit cell, there is the equivalent of four shared atoms.

For a silicon crystal, the unit cell is a variation of the FCC known as the *FCC diamond structure* (see Figure 4.6). Atoms are shared as with the FCC unit cell, plus there are four complete atoms situated inside the cubic structure. For a silicon unit cell, there are a total of eight complete atoms, with four shared and four unshared.

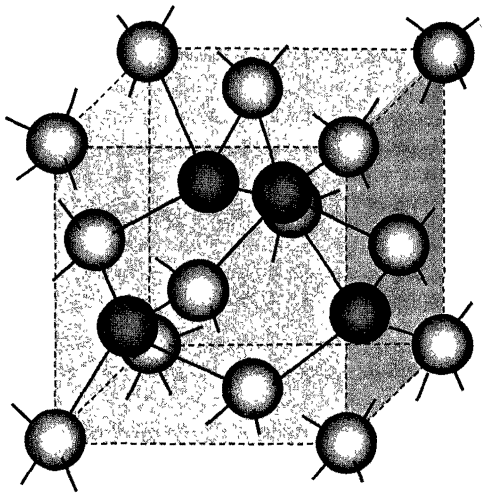


FIGURE 4.6 Silicon Unit Cell: FCC Diamond Structure

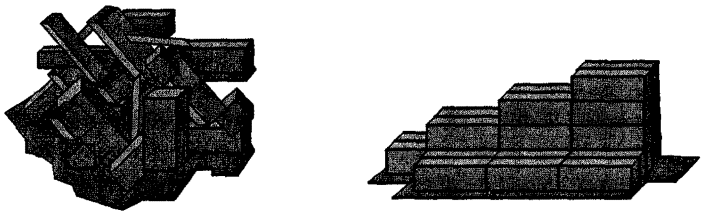
Polycrystal and Monocrystal Structures

Now that the unit cell is defined for a crystal structure, we can define how the unit cells are organized. If the unit cells are not in a regular arrangement, then the material is a *polycrystal.* The

silicon produced from the semiconductor-grade purification process is polycrystalline silicon. Another name for polycrystalline silicon material is *polysilicon.* An analogy for a polycrystal structure is a pile of bricks. Each individual brick represents the unit cell and the pile is polycrystalline because there is no repetitive arrangement (see Figure 4.7).

If the unit cells are neatly arranged in a three-dimensional, repeatable manner, then the crystal structure is *monocrystal.* Another term often used for monocrystal is *single crystal.* For our brick analogy, if the bricks are now stacked neatly in repeating rows next to one another, then the bricks represent a monocrystal. A monocrystal structure is also shown in Figure 4.7.

Semiconductor wafer processing requires a pure, monocrystalline silicon structure (single crystal). This is because the repeatable unit cells of a monocrystal structure provide the desirable electrical and mechanical properties necessary for silicon wafer processing and performance. Unacceptable crystal structure and defects influence the formation of microdefects that affect wafer processing (see the following section on crystal defects in silicon).



Polycrystalline structure

Monocrystalline structure

FIGURE 4.7 Polycrystalline and Monocrystalline Structures

CRYSTAL ORIENTATION

Before studying the production process to obtain a monocrystal structure for silicon, let's first examine how unit cells are oriented in silicon. Unit cell orientation is important because it defines how the silicon crystal structure is physically aligned with the silicon wafer. Different crystal orientations change the chemical, electrical, and mechanical properties of the silicon, which affects the manufacturing process conditions and final device performance.

To define the silicon unit cell orientation, we need a coordinate system. In crystals, the coordinate system has three axes, *x, y,* and *z,* shown in Figure 4.8. We will arbitrarily assign the number 1 at an equal distance along each axis, while the center point is given the value of 0. These are referred to as unit values. If the crystal is a monocrystal, then all unit cells are aligned in a repetitive manner along these three coordinate axes.

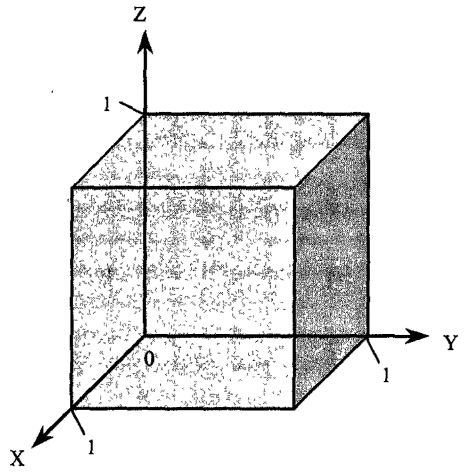


FIGURE 4.8 Axes of Orientation for Unit Cells

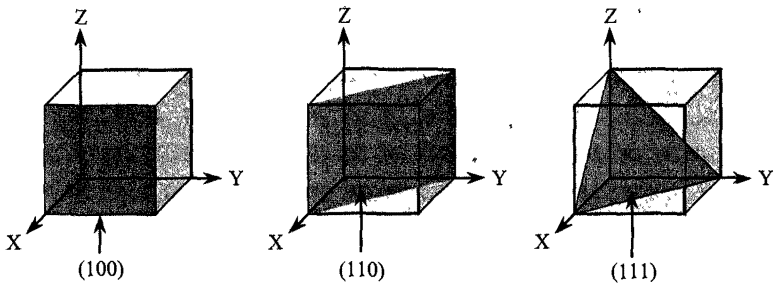


FIGURE 4.9 Miller Indices of Crystal Planes

The orientation of planes in the silicon crystal is described by a set of numbers known as *Miller indices.*[6] It is beyond the scope of this book to completely describe the Miller notation system for crystal planes and their directions. In the Miller system of notation, parentheses ( ) are used to denote a specific plane, whereas brackets < > denote groups of equivalent directions.[7]

From our standpoint, what is important is to recognize that each set of Miller indices enclosed in a parenthesis specifies a unique plane in the crystal structure. For semiconductor fabrication, the Miller notations applicable to the most common crystal planes on a wafer are (100), (110), and (111). These three crystal planes are shown in Figure 4.9 on page 72. They are obtained in the silicon crystal by maintaining precise control of orientation during the crystal growth process (described in the next section). You interpret each Miller notation based on where the plane intersects the axes.

The (100) crystal plane is parallel to the y-z plane and intersects the x-axis at the unit value of one. The (110) intersects the x- and y-axes only, whereas the (111) intersects the x-, y-, and z-axes. Wafers with a (100) crystal plane orientation are most common for fabricating MOS devices. The reason for this is that the surface state condition for (100) silicon is more conducive toward controlling the threshold voltage required to turn MOS devices on and off.[8] The (111) crystal plane orientation has a tighter packing density at the atomic level, making it easier to grow. These are the least expensive crystals to grow and are often used for bipolar devices.[9] Gallium arsenide (GaAs) technology also uses wafers with (100)-oriented planes.

MONOCRYSTAL SILICON GROWTH

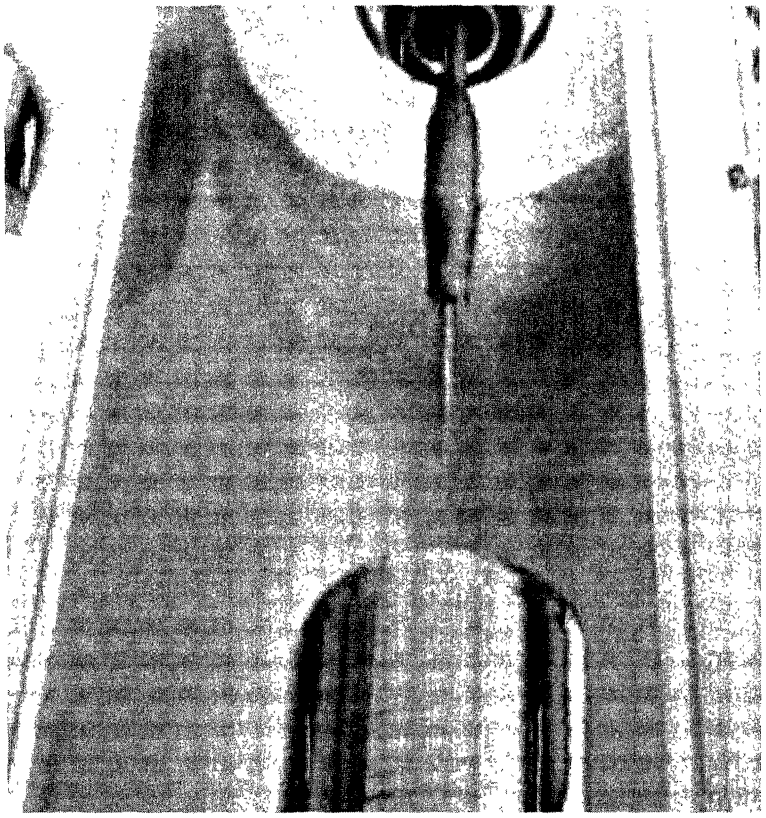
Crystal growth is the process of converting the polysilicon chunks of semiconductor-grade silicon into a large monocrystal of silicon. The grown silicon monocrystal is referred to as an *ingot.* The most common technique today for growing a silicon monocrystal ingot to be used in wafer fabrication is the *Czochralski (CZ) method,* which is named for the inventor who developed the process in the early 1900s.

CZ Method

The Czochralski (CZ) growth of a silicon monocrystal involves the transformation of molten SG silicon liquid into a solid silicon ingot that has the correct crystal orientation and is doped either as n-type or p-type. Over 85% of all silicon crystals are grown according to the CZ method.[10]

A piece of monocrystalline silicon having the desired crystal orientation is used as a starter seed to grow the silicon ingot that is a replica of the original seed crystal. To achieve crystal growth with the CZ method, precise conditions are controlled at the contact interface between molten silicon and a single monocrystal seed of silicon. These conditions permit thin films of silicon to accurately replicate the seed crystal structure and grow into a large ingot. This is done with equipment known as a *CZ crystal puller.*

CZ Crystal Puller ■ To grow a silicon ingot, chunks of SG silicon are placed in a fused silica (amorphous quartz) crucible holder, along with small amounts of dopant to create either n-type or p-type silicon. The crucible is large, with crucible diameters for a 300-mm wafer at 32 inches or larger. These crucibles must hold from 150 to 300 kg of silicon.[11] An alternative approach for larger diameter wafers is to use granular polysilicon in the crucible, which permits the gradual



Silicon Ingot Grown by CZ Method
(Photo courtesy of Kayex Corporation)

introduction of the silicon during the meltdown to minimize stresses on the large crucible during the melting process.[12] The crucible is positioned in the crystal puller, which is where the silicon crystal ingot is grown (see Figure 4.10).

The crucible of silicon is heated in the furnace of the puller using either resistance heaters or RF (radio frequency) heating coils. Resistance heaters are used for larger diameter ingots. As the silicon is heated, it turns to liquid and is referred to as the *melt*. A seed of a perfect silicon crystal is attached to a pull mechanism to start the structure of the new crystal. The seed is placed at the surface of the melt and is slowly drawn away while rotating in a direction that is opposite to the rotation of the crucible.

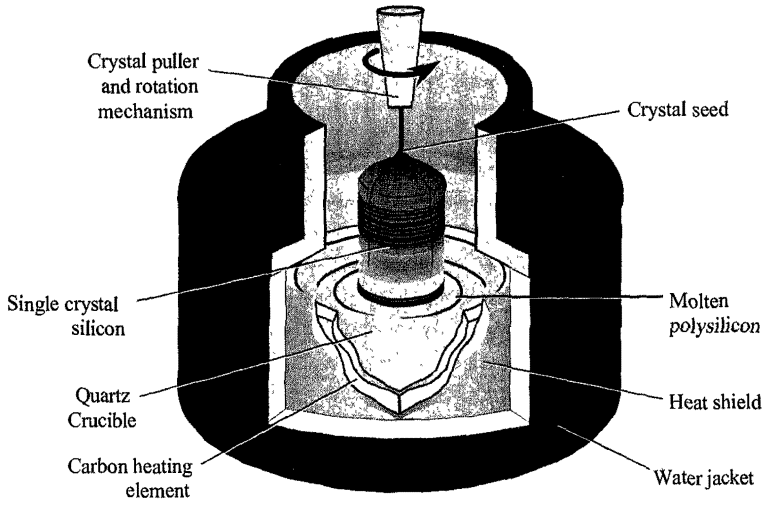
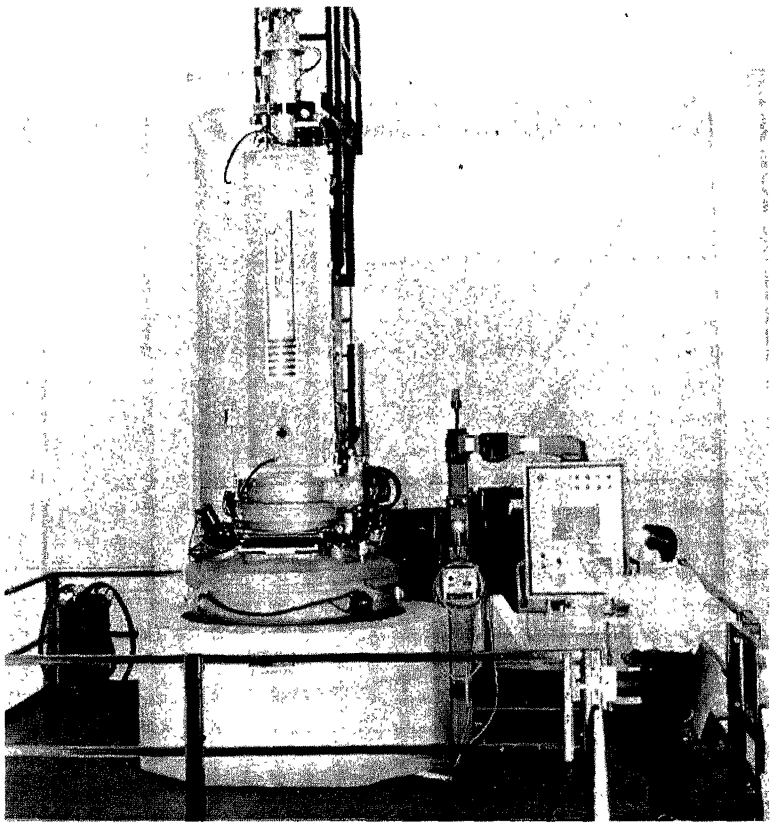


FIGURE 4.10 CZ Crystal Puller



CZ Crystal Puller
(Photo courtesy of Kayex Corporation)

As the single seed is drawn away from the melt during the pull process, liquid from the melt is raised with it by surface tension. The interface at the single seed dissipates heat and solidifies downward toward the melt. As the seed is steadily pulled out of the melt and rotated, a single crystal grows with the same crystal orientation as the seed. While the seed is rotated, the crucible is also rotated. Different ingot growth patterns result depending on the speed of the seed and crucible rotation and their respective directions.

The objectve of the pull process is to precisely replicate the seed structure while obtaining dopant uniformity, achieving the correct ingot diameter, and limiting introduction of impurities into the silicon. The two main parameters that affect the pull process are the pull rate and crystal rotation.[13] The pull starts at a relatively high speed and then slows down considerably. This action forms a neck in the ingot, since the diameter of the growing crystal is directly related to the rate of pull.

A goal of the crystal growth process is to have uniform (or homogeneous), large-diameter crystals. This result gives more predictable crystal properties with respect to parameters such as resistivity. A technique developed to achieve this with the CZ process is to use a magnetic field around the silicon melt to stabilize the crystal during its growth.[14] This condition is referred to as *magnetic CZ (MCZ)*.

Doping ■ Dopant material is added to the melt in the crystal puller in order to obtain the desired electrical resistivity in the finished crystal. The resistivity of pure silicon is approximately $2.5 \times 10^5\ \Omega$-cm. The most common dopants used in crystal growth are trivalent boron to create p-type silicon or pentavalent phosphorus to create n-type silicon. Dopant concentration ranges in the silicon are indicated through letter designations and superscripts, as shown in Table 4.2. Usually raw dopants are not added directly to the melt because the dopant amounts are extremely small. It is typical to add the dopant in the form of a highly doped powder of crushed silicon.[15]

TABLE 4.2 Dopant Concentration Nomenclature in Silicon

Dopant	Material Type	Concentration (Atoms/cm^3) $< 10^{14}$ (Very Lightly Doped)	10^{14} to 10^{16} (Lightly Doped)	10^{16} to 10^{19} (Doped)	$>10^{19}$ (Heavily Doped)
Pentavalent	n	n^-	n^-	n	n^+
Trivalent	p	p^-	p^-	p	p^+

Impurity Control ■ Impurity control is important during crystal growth since unacceptable impurities affect the device performance. One impurity that must be controlled but that can also be beneficial is oxygen. The primary source of oxygen in the CZ method results from the dissolution of the crucible during crystal growth.[16]

Small amounts of oxygen in the ingot are desirable because oxygen can serve as an internal getter to tie up metallic contaminants that are introduced into the wafer during the fabrication process. Note that the term *getter* is used to describe any process that immobilizes or ties up impurities. Most of the oxygen in the silicon crystal near the wafer surface is removed during the many heating processes that take place during wafer fabrication, leaving higher oxygen concentrations deeper in the wafer. This oxygen then serves as the getter that attracts the contaminants away from the devices located on the surface. The amount of oxygen required to serve as a getter is continually being reduced as the device fabrication process is improved to reduce the sources of contaminants.[17]

Float-Zone Method

An alternative crystal growth method is the *float-zone* method, which produces a silicon monocrystal ingot with significantly lower oxygen content.[18] The float-zone method was developed in the 1950s and produces the purest bulk silicon single crystals known to date. A schematic of the float-zone process is shown in Figure 4.11.

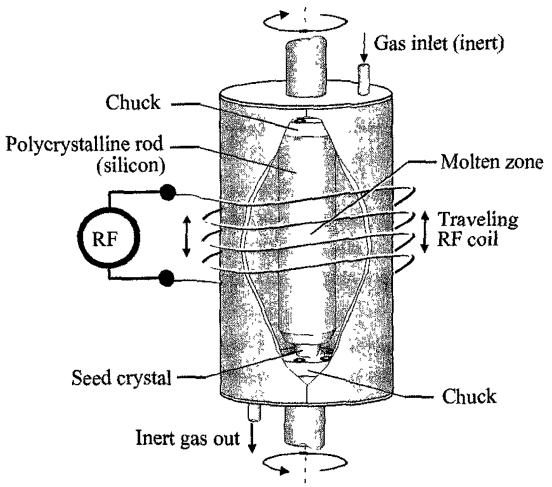


FIGURE 4.11 Float Zone Crystal Growth

The float-zone method of growing a monocrystal ingot of silicon starts with a bar of doped polysilicon that has been cast in a mold. A monocrystal seed is attached to one end of the bar and then placed in the grower. An RF (radio frequency) coil applies heat at the contact area of the bar and seed. Heating of the polysilicon bar is the most important aspect of the float-zone process, as each section of the bar is molten only for about 30 minutes before it solidifies again at the monocrystal interface.[19] This monocrystal growth process progresses along the bar as the heater is moved along its axis.

The float-zone process typically makes smaller-diameter wafers than the CZ process and was producing predominantly 125-mm wafers at the turn of the century. By not using a crucible, the float-zone process results in high-purity silicon with lower oxygen content.

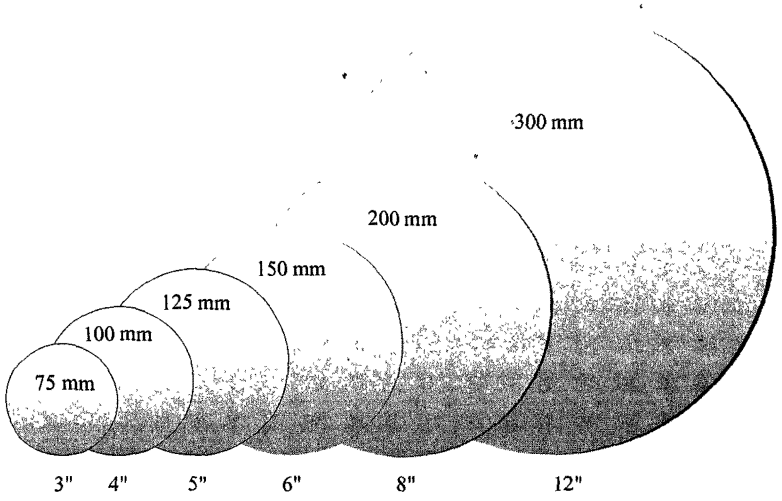


FIGURE 4.12 Wafer Diameter Trends

Reasons For Larger Ingot Diameters

Silicon ingot diameters have increased from the early diameters of less than 25 mm in the 1950s to the most recent diameter of 300 mm. The historical trend for larger semiconductor wafer diameters is shown in Figure 4.12.

There are wafer fabrication facilities still in use that produce wafers with diameters of 75 mm, 100 mm, 125 mm, and 150 mm. The most common practice is to build a new fab to introduce a new wafer diameter, since upgrading a facility to a larger wafer diameter can cost hundreds of millions of dollars. The semiconductor industry began to transition to 300-mm wafers around the year 2000. There are also early evaluations underway to understand the possibility of increasing wafer diameters to 400 mm. Table 4.3 highlights different attributes of various wafer sizes.

TABLE 4.3 Wafer Dimensions & Attributes

Diameter (mm)	Thickness (μm)	Area (cm^2)	Weight (grams/lbs)	Weight/25 Wafers (lbs)
150	675 ± 20	176.71	28 / 0.06	1.5
200	725 ± 20	314.16	53.08 / 0.12	3
300	775 ± 20	706.86	127.64 / 0.28	7
400	825 ± 20	1256.64	241.56 / 0.53	13

From H. Huff, R. Goodall, R. Nilson, and S. Griffiths, "Thermal Processing Issues for 300 mm Silicon Wafers: Challenges and Opportunities," *ULSI Science and Techonology* (New Jersey: The Electrochemical Society, 1997), p. 139.

Larger ingot diameters bring new challenges to achieving correct crystal growth and maintaining correct process control during ingot growth. Ingots with a 300-mm diameter are about 1 meter long and require between 150 to 300 kg of SG silicon in the crucible for the melt. With the increased complexity needed to manufacture the silicon ingot, what is the main reason to keep increasing the diameter? The answer lies in the cost benefits to the wafer fabrication process due to increasing the wafer diameter.

Larger-diameter wafers have a greater surface area for chips. For 300-mm wafers, there is 2.25 times the surface area than that available on a 200-mm wafer (see Figure 4.13). This increase translates into more chips that can be produced from a single wafer.

An analogy would be driving a car 200 miles from city A to city B. If the driver travels alone, then this establishes a certain cost for making the trip (e.g., fuel, wear on the vehicle, and so on). However, if a passenger goes with the driver, then the driver and the passenger make the trip for less cost. If three passengers make the trip with the driver, then there are even more savings per passenger. This approach to efficiency is called *economies of scale.*

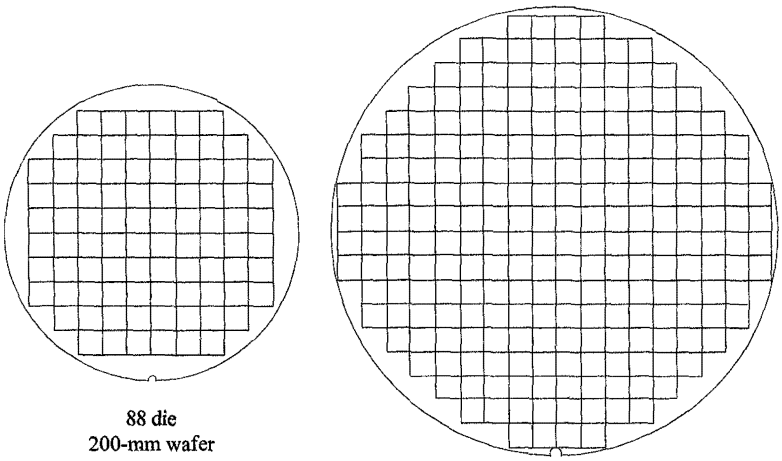


FIGURE 4.13 Increase in Number of Chips on Larger Wafer Diameter

Larger wafers means more chips per wafer, leading to improved equipment productivity due to economies of scale for less handling per chip and less processing time per chip. It is estimated that conversion to a 300-mm wafer diameter will reduce the fabrication cost per chip by approximately 30% through equipment utilization improvements.[20] Another benefit from larger-diameter wafers is that fewer chips are near the edge of the wafer, which translates into higher production yield. There is also a benefit from tool repeatability from chip to chip since more chips are exposed to the same process conditions.

The total industry cost to convert the standard wafer diameter from 200 mm to 300 mm has been estimated at $13 billion to $15 billion. No one firm or country has the resources to make this change alone. Global 300-mm standards have been developed and 23 semiconductor companies worldwide have joined two separate consortia to coordinate the conversion to 300-mm wafers. The two consortia are the International 300 mm Initiative (I300I), which includes the U.S., Europe, Korea and Taiwan, and the Semiconductor Leading Edge Technology (Selete), which represents ten IC companies from Japan.[21] Ultimately, cost savings will be the major factor that drives the industry toward a larger wafer diameter. A summary of dimensional specifications for 300-mm wafers is provided in Table 4.4.

TABLE 4.4 Developmental Specifications for 300-mm Wafer Dimensions and Orientation Requirements

Parameter	Units	Nominal	Some Typical Tolerances
Diameter	mm	300.00	± 0.20
Thickness (center point)	μm	775	± 25
Warp (max)	μm	100	
Nine-Point Thickness Variation (max)	μm	10	
Notch Depth	mm	1.00	+ 0.25, -0.00
Notch Angle	Degree	90	+5, -1
Back Surface Finish		Bright Etched/Polished	
Edge Profile Surface Finish		Polished	
FQA (Fixed Quality Area—radius permitted on the wafer surface)	mm	147	

From H. Huff, R. Goodall, R. Nilson, and S. Griffiths, "Thermal Processing Issues for 300 mm Silicon Wafers: Challenges and Opportunities," *ULSI Science and Technology* (New Jersey: The Electrochemical Society, 1997), p. 139.

CRYSTAL DEFECTS IN SILICON

Semiconductors require a nearly perfect silicon crystal structure in order to function optimally for advanced ICs. A *crystal defect* is any interruption in the repetitive nature of the unit cell crystal structure. Another term used for a crystal defect is *microdefect.* It is essentially impossible to grow or process silicon without making a defect. However, progress has been made on growing nearly defect-free silicon with a very low defect density.

Defect density is a commonly used term in crystal growth and wafer fabrication. It is defined as the number of defects per cm^2 of wafer surface that may occur during processing due to all kinds of causes. Reducing defect density is a critical aspect for increasing wafer yield (see Figure 4.14). Yield is the percentage of good chips produced out of the total chips on a wafer.

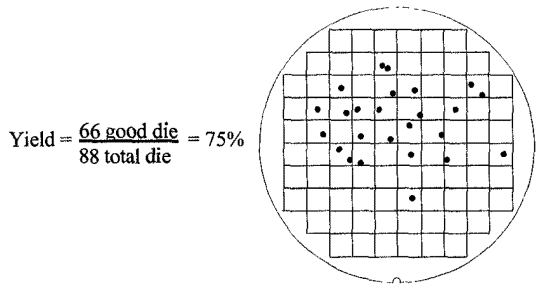


FIGURE 4.14 Yield of a Wafer

It is important to study silicon crystal defects because they have a damaging effect on the electrical properties of the semiconductor. These problems includes SiO_2 dielectric breakdown and leakage current failures. As device geometries shrink and active gate areas increase due to more transistors on a chip, the probability of a defect occurring in a sensitive area of the chip increases. Such a defect can negatively affect the device yield for advanced ICs.

Crystal defects can occur during crystal growth and the subsequent processing of the silicon ingot and wafers. Some crystal defects occur due to surface damage, including cracks and surface defects from mechanical operations on the wafer. There are three general forms of crystal defects in silicon:

1. Point defects: Localized crystal defect at the atomic level
2. Dislocations: Displaced unit cells
3. Gross defects: Defects in crystal structure

Point Defects

A *point defect* occurs at a particular location in the crystal lattice. Three types of point defects are shown in Figure 4.15. The most basic type of point defect is a *vacancy* (also referred to as a *void*).

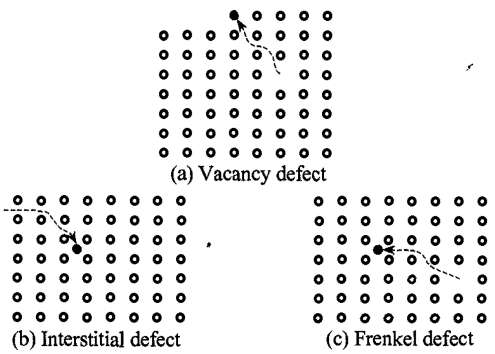


FIGURE 4.15 Point Defects
Redrawn from S. Ghandi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide,* 2nd ed. (New York: Wiley, 1994), p. 23.

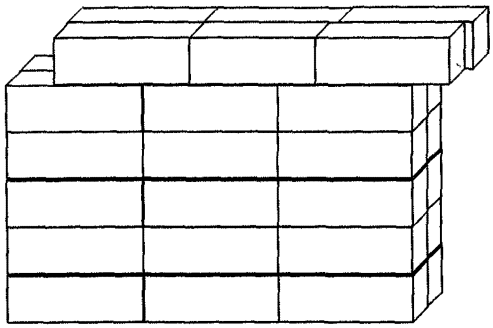


FIGURE 4.16 Dislocations

This defect occurs when an atom is removed from its lattice site and moves to the surface of the crystal. Another type of point defect is an *interstitial,* which occurs when an atom becomes located in one of the many voids within the crystal structure. When an atom leaves its lattice site and takes up position in a void, it creates a vacancy-interstitial pair, or *Frenkel defect.*

Point defects in semiconductor silicon have become increasingly important as device technology becomes more complex. Crystal growth conditions that affect the creation of point defects during wafer growth are the growth rate (how fast the crystal is pulled) and the thermal gradient (the temperature difference between the melt and the solid crystal) at the crystal-melt interface.[22] The silicon crystal cooling rate is controlled, which effectively reduces the defect generation. High-temperature processing during semiconductor manufacturing is also a source of point defects.

Another type of point defect is due to chemical impurities that are introduced into the lattice. Impurities are introduced intentionally or unintentionally during fabrication. They can take up locations at sites ordinarily occupied by atoms, known as *substitutional impurities,* or locate themselves in voids as interstitial impurities.

The most significant impurities not introduced intentionally into the silicon are oxygen and carbon. It is important in VLSI and ULSI device fabrication to control these impurities. Oxygen enters the silicon melt via dissolution of the crucible during the CZ process, as discussed earlier in this chapter. Most of this oxygen is removed, but less than 5% remains in the growing crystal. Oxygen present in silicon can create crystal defects such as dislocations (see the next section). The presence of donor electrons from oxygen impurities in the silicon lattice can also affect the electrical parameters of a pn junction. Carbon can act as a nucleation site for crystalline defect formation in silicon.

Dislocations

Unit cells form a repetitive structure in a monocrystal. If unit cells are displaced, this condition is known as a *dislocation* (see Figure 4.16). An analogy would be a neat stack of bricks that has a group of bricks displaced along a row. A form of dislocation is *stacking faults* which is due to layer stacking errors. Dislocations can be introduced at any stage of the crystal growth or wafer fabrication process. However, dislocations occurring after crystal growth are often associated with mechanical stress to the wafer, such as that caused by uneven heating and cooling or excessive force applied to the wafer.

In some instances, dislocation faults are induced following the thermal oxidation of the silicon wafer surface during device fabrication (see Chapter 10).[23] The result is referred to as *oxygen-induced stacking faults* (*OISF*). OISF defects appear as shallow, saucer-shaped depressions in the lattice. These defects can be detected through X-ray analysis or surface etching. Specialized thermal treatments (referred to as annealing) or gettering can be used to minimize stacking faults and dislocations in silicon crystal structures.

Crystal growth dislocations have been reduced for large-diameter wafers through use of a neck-down procedure during ingot growth.[24] This procedure consists of necking down the cross section of the single seed at the beginning of the pull so that the dislocations grow out to the surface, and then growing several centimeters at sufficiently high speed so that high-vacancy densities remove the edge dislocations. However, as described above, dislocations can still occur from other

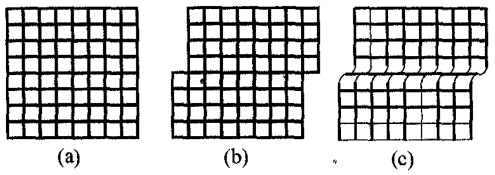


FIGURE 4.17 Crystal Slip
Redrawn from S. Ghandi. *VLSI Fabrication Principles: Silicon and Gallium Arsenide,* 2nd ed. (New York: Wiley, 1994), p. 49.

sources during the wafer fabrication process, such as from wafer edge chipping or high temperature processing.

Gross Defects

Gross defects are related to the structure of the crystal and typically occur during crystal growth. A *slip* is a gross defect in a crystal that occurs when there is slippage of the crystal along one or more crystal planes (see Figure 4.17).

Another gross defect is *twin planes* (or *twinning*), where the crystal grows in two different directions from the same plane (see Figure 4.18). The source of twin planes could be thermal or mechanical shock during the growth process. The crystal on either side of the defect may be perfect. Crystals with either slip or twin planes are unacceptable for semiconductor manufacturing.

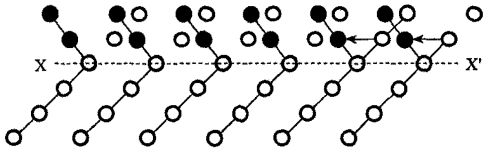


FIGURE 4.18 Crystal Twin Planes
Redrawn from S. Ghandi. *VLSI Fabrication Principles: Silicon and Gallium Arsenide,* 2nd ed. (New York: Wiley, 1994), p. 55.

WAFER PREPARATION

Silicon is a hard, brittle material and of little use for semiconductor manufacturing in its form as an ingot after crystal growth. The cylindrical, single-crystal ingot (also called a *boule*) must undergo a series of process steps to transform it into wafers that meet stringent specifications for semiconductor manufacturing. These wafer preparation steps include machining operations, chemical operations, surface polishing, and quality measures. The basic process flow for wafer preparation is shown in Figure 4.19.

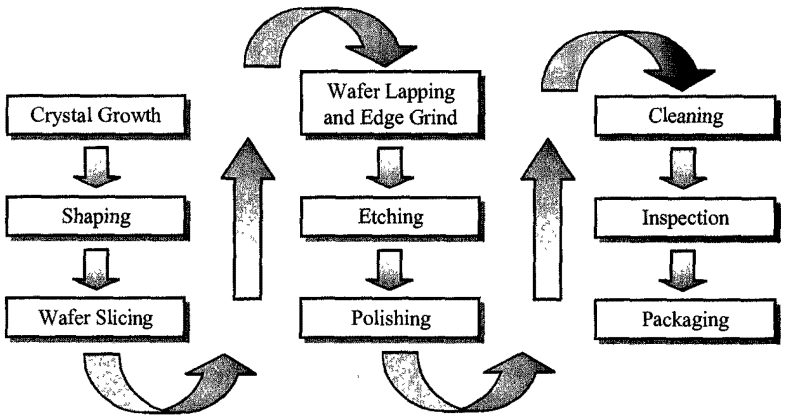


FIGURE 4.19 Basic Process Steps for Wafer Preparation

Because of the continued advances in chip design and fabrication requirements, the wafer preparation process must be capable of delivering wafers that meet tighter specification requirements. Wafer preparation requirements include the wafer geometry (diameter, flatness, and warpage), surface perfection (roughness and light scattering), and cleanliness (sources of particles). These wafer specifications address issues such as dimensional control for robotic material handling or surface conditions required for critical process steps during IC fabrication.

Shaping Operations

After growth of the ingot in the crystal puller, the shaping operations are the first process steps. Shaping operations include all preliminary steps to prepare the monocrystal silicon ingot prior to slicing into wafers.

End Removal ■ The first operation is to remove the two ends of the ingot. The ends are referred to as the *seed end* (where the seed was located) and the *tang end* (opposite to the seed end). Once the ends are removed (cropped off), a four-point probe resistivity check is done to confirm the proper dopant uniformity throughout the ingot (see Chapter 7 for the resistivity measurement procedure).

Diameter Grinding ■ The next operation is *diameter grinding* to create the precise diameter of the material. Ingots are grown slightly oversize to permit this grinding step, since diameter control and roundness cannot be adequately controlled during crystal growth. Precise diameter control is critical for semiconductor manufacturing given the automated wafer-handling steps in wafer fabs. Figure 4.20 illustrates the diameter-grinding process.

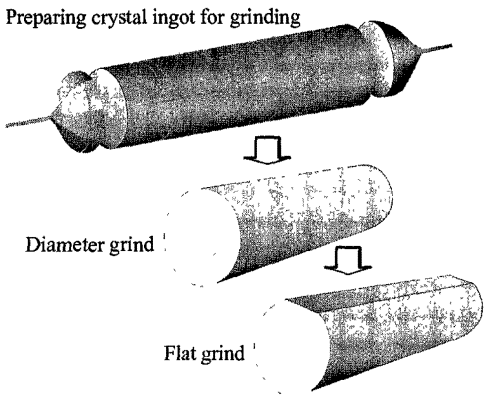


FIGURE 4.20 Ingot Diameter Grind

Wafer Flat or Notch ■ The semiconductor industry has traditionally placed *flats* on the ingot to identify the crystal structure and to orient the wafer. The primary flat orients the wafer to the crystal structure, as shown in Figure 4.21 on page 82. There is a secondary flat which identifies the orientation and conductivity type of the wafer.

Wafer flats have been replaced in the U.S. with a *notch* for wafers 200 mm and larger. Notched wafers have information about the wafer laser-scribed in a small region on the wafer. The location and depth of the laser-scribing initially created some fear of contamination collecting in the small laser marks. For 300-mm wafers, a standard has been approved for laser-scribing all wafers on the back surface in an exclusion zone near the wafer's edge.[25] The notch and laser scribe are illustrated in Figure 4.22 on page 82.

For 300-mm wafers, the exclusion zone is outside an area known as the fixed-quality area (FQA), which is where chips are permissible on the wafer surface. The exclusion zone is currently 3 mm, but may be reduced to 2 mm at some time in the future.

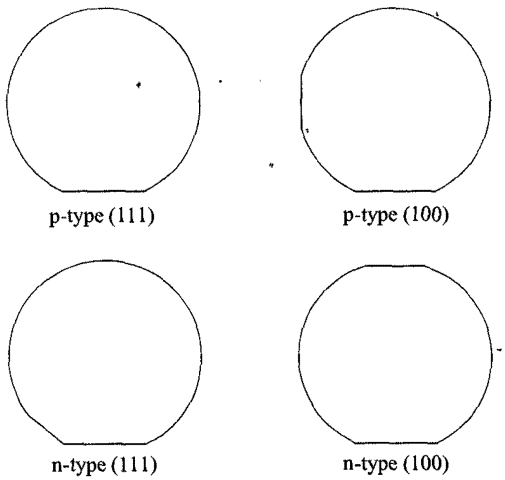


FIGURE 4.21 Wafer Identifying Flats

Wafer Slicing

Once the shaping operations are complete, the ingot is ready to undergo *wafer slicing*. This is the first major process step after ingot growth. For wafers up to 200 mm in diameter, slicing is done with an internal diameter saw that has a diamond cutting edge, as seen in Figure 4.23. The internal diameter saw is used because the cutting edge is more stable to yield a flat cutting surface.

For 300-mm wafers, internal diameter saws are not desirable for slicing wafers due to the larger diameter. Wafer slicing for 300-mm ingots is currently being done with wire saws.[26] Wire saws yield more wafer slices per inch of crystal than conventional ID saws due to thinner kerf (saw blade thickness) losses associated with using a slurry-coated wire instead of a diamond-coated blade.[27] Wafer sawing reduces mechanical damage to the wafer surface during the slicing process, but there are still concerns about inadequate wafer-flatness control with sawing.[28] Equipment development is ongoing for wafer slicing of 300-mm and larger wafers to obtain precise wafer dimensions.

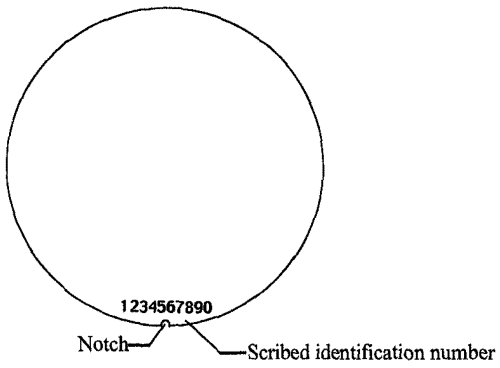


FIGURE 4.22 Wafer Notch and Laser Scribe

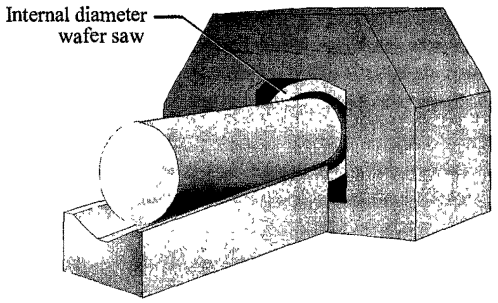


FIGURE 4.23 Internal Diameter Saw

The thickness of the wafer is carefully controlled during the slicing process. The 300-mm wafer is currently specified at a thickness of 775 ± 25 microns. Thicker wafers are better able to withstand thermal and mechanical shocks introduced during high temperature processing in the semiconductor fabrication process.

Wafer Lapping and Edge Contour

Following wafer slice, there is traditionally a mechanical, two-sided *lapping* operation to remove damage left by slicing and to achieve a high degree of parallelism and flatness of the wafer. Lapping is performed under rotational pressure with pads and an abrasive slurry mixture that typically consists

of alumina or silicon carbide and glycerine. Flatness is a critical wafer parameter for many process steps in the wafer fabrication process.

A *polished wafer edge* finish (also called *edge grind*) is applied to the wafer to contour a smooth radius on the edge of the wafer (see Figure 4.24). This step may takes place either before or after lapping. Cracks and small crevices at the edge of the wafer create mechanical stress in the wafer that activates crystal dislocations, especially during thermal process steps that occur in wafer fabrication. Crevices can also be the source of unwanted contamination buildup as well as particulate flaking during fabrication. A smooth edge radius is important to minimize these concerns.[29] Furthermore, chipped edges are a source of edge dislocation growth during thermal cycles in the wafer fabrication process.

Etching

Shaping the wafer leaves the surface and edges damaged and contaminated. The depth of the wafer damage depends on the particular process at a manufacturer, but it is usually on the order of several microns deep. To remove the damaged surface from the wafer, wafer suppliers use a technique known as wafer etch or chemical etching. *Wafer etch* is a process involving the chemical removal of selective surfaces of a material (see Figure 4.25). Wafers undergo a wet chemical etch process to remove the damage and contamination on the wafer. Typically about 20 microns of silicon wafer surface is removed during the etch process to ensure that all damage is removed.[30] The etching is done with either acid or alkaline chemicals, depending on the process in place at the wafer manufacturer. The subject of etching will be covered in detail in Chapter 16.

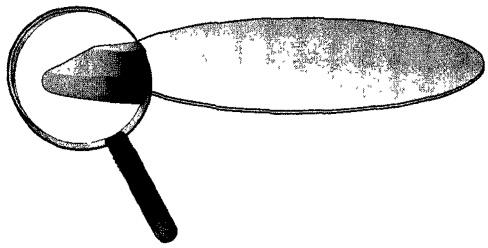


FIGURE 4.24 Polished Wafer Edge

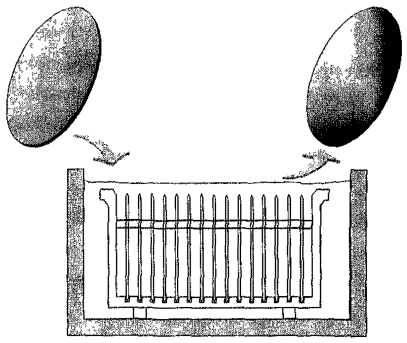


FIGURE 4.25 Chemical Etch of Wafer Surface to Remove Damage
Redrawn from M. S. Bawa, E. F. Petro, H. m. Grimes, "Fracture Strength of Large Diameter Silicon Wafers," *Semiconductor International* (November 1995), p. 115.

Polishing

The final step in the wafer preparation is a *chemical mechanical planarization* (*CMP*) to achieve a smooth wafer surface with a high degree of flatness. CMP is also referred to as polish (see Chapter 18 for a discussion of CMP). This CMP step has traditionally been done only on the topside of the wafer for 200-mm and earlier wafers, leaving the chemically-etched surface on the backside. This etched backside left a relatively rough surface, being approximately three times rougher then a surface processed with CMP.[31] Its purpose was to give a rougher surface for handling devices. However, there has been concern about the ability of an etched surface to meet the flatness requirements for deep submicron photolithography, plus the possibility of introducing particulate contamination into the wafer fabrication process.

For 300-mm wafers, double-sided polishing (DSP) with CMP is the final major manufacturing step. The planetary motion of the wafer between the polishing plates produces flat and parallel surfaces while improving surface roughness (see Figure 4.26 on page 84). Since this is the final wafer preparation step, flatness across the large wafer diameter can be readily maintained. The polished back surface also permits the wafer manufacturer to characterize wafer cleanliness prior to delivery to the wafer fab. The final wafer surface has a mirrorlike finish on both sides of the wafer.

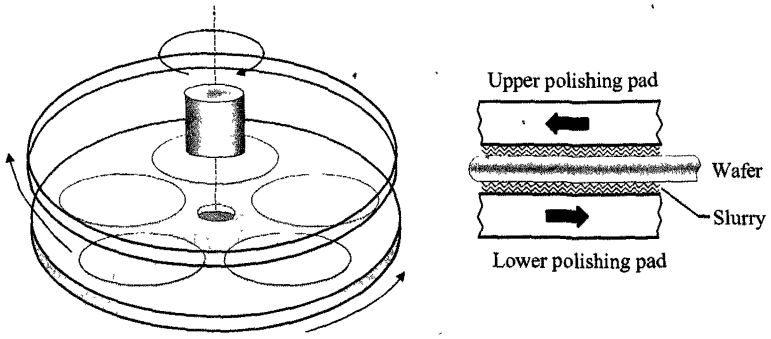


FIGURE 4.26 Double-Sided Wafer Polish

Cleaning

Semiconductor wafers must be cleaned to achieve an ultraclean state of cleanliness prior to shipment to wafer fabrication facilities. Cleaning specifications have undergone considerable development in the last few years, to the extent that wafers should be free of particles and contamination. The different types of wafer contamination, the appropriate cleaning steps, and the process procedures for cleaning are discussed in Chapter 6.

Wafer Evaluation

Prior to packaging the wafer, it is inspected for quality criteria that are specified by the customer. Standard quality measures are discussed later in this chapter. The most critical criteria relate to surface defects such as particulate contamination and stains.

Packaging

Wafer suppliers must carefully package wafers for shipment to the wafer fabrication facilities. Considerable effort can be lost if the wafers are damaged from transit or from the packaging material. The wafers are stacked in a plastic cassette or "boat" that has grooved slots to hold the wafers. A fluorocarbon resin material such as Teflon is often used as the cassette material to minimize particle generation. In addition, the Teflon is made conductive so that it does not generate an electrostatic discharge. All equipment and operators are grounded to drain off any charge buildup that could attract particles.

Once filled with wafers, the cassette is then placed in a nitrogen-filled, sealed canister that prevents oxidation and other contamination during transit. When the wafers are received at the wafer fab, they are transferred to other standardized cassettes used for transport and handling during wafer processing inside the fabrication facility. The shipping container is designed to minimize the need for wafer handling. One shipping concept under development is a 25-wafer container referred to as a front-opening shipping box (FOSB) which would interface with automated material handling systems in the wafer fab.[32]

QUALITY MEASURES

Silicon wafer suppliers produce wafers to requirements that are becoming more stringent, which represents improved control over wafer quality. A sampling of some wafer specifications and their rate of improvement relative to the critical dimension on a wafer is shown in Table 4.5.

For wafer quality measures, uniformity throughout the wafer is critical. Important silicon wafer quality requirements are:

- Physical dimensions
- Flatness
- Microroughness
- Oxygen content
- Crystal defects
- Particles
- Bulk resistivity

The suppliers of silicon wafers must control their wafer quality by performing ingot and wafer quality inspections to demonstrate that quality specifications are met. Defective wafers shipped unknowingly to a semiconductor manufacturing wafer fab facility would be catastrophic.

TABLE 4.5 Improving Silicon Wafer Requirements

	Year (Critical Dimension)			
	1995 (0.35 μm)	**1998 (0.25 μm)**	**2000 (0.18 μm)**	**2004 (0.13 μm)**
Wafer diameter (mm)	200	200	300	300
Site flatness[A] (μm)	0.23	0.17	0.12	0.08
Site size (mm × mm)	(22 × 22)	(26 × 32)	26 × 32	26 × 36
Microroughness[B] of front surface (RMS)[C] (nm)	0.2	0.15	0.1	0.1
Oxygen content (ppm)[D]	≤ 24 ± 2	≤ 23 ± 2	≤ 23 ± 1.5	≤ 22 ± 1.5
Bulk microdefects[E] (defects/cm^2)	≤ 5000	≤ 1000	≤ 500	≤ 100
Particles per unit area (#/cm^2)	0.17	0.13	0.075	0.055
Epilayer[F] thickness (± % uniformity) (μm)	3.0 (± 5%)	2.0 (± 3%)	1.4 (± 2%)	1.0 (± 2%)

Adapted from K.-M. Kim, "Bigger and Better CZ Silicon Crystals," *Solid State Technology* (November 1996), p. 71.

Notes:
A: Flatness is the linear thickness variation across the wafer or a site on a wafer (see the following).
B: See the following for a description of microroughness.
C: RMS is a method for determining the best estimate of group of measurements—in this case, the surface finish measurements (see the following). It is calculated by taking the root-mean-square (square root of the average of all measurements squared). Surface finish measurements are obtained by measuring the highest point relative to the lowest point on a surface.
D: ppm is part per million.
E: Bulk microdefects represent all defects within a square centimeter.
F: See the following section for a description of epilayer.

Physical Dimensions

Wafers must be physically dimensioned in order to meet the requirements for device fabrication and to accommodate the automated wafer handling equipment used in the wafer fab. The wafer physical dimensions controlled and inspected during wafer preparation include measurements such as diameter, thickness, location and size of orientation, flat (or notch), and wafer deformation. Figure 4.27 shows an example of wafer deformation. The most likely source of wafer deformation is the slicing process.

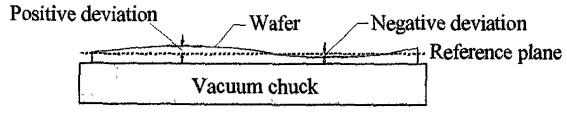


FIGURE 4.27 Wafer Deformation

Flatness

Flatness is one of the most critical wafer parameters specified, primarily because the process of photolithography (Chapters 13 to 15) is highly sensitive to local-site flatness. *Wafer flatness* is the linear thickness variation across the wafer. It is measured by determining the distance that the front surface of the wafer deviates from a specified reference plane. For a wafer, the reference plane is theoretically defined by the back wafer surface if it were held perfectly flat, such as when pulled down by a vacuum onto a clean, flat surface (see Figure 4.28). Note that the front surface of the wafer is the wafer side where chips will be placed.

Flatness can be specified as site flatness for a specific site on the wafer or specified as global flatness, which is the total wafer flatness across the fixed-quality area (FQA) of the-wafer surface. The FQA does not include the exclusion zone on the periphery of the wafer surface. A larger area for measuring flatness is more difficult to control than a smaller area. Typical flatness specifications for site flatness are shown in Table 4.5 on page 85.

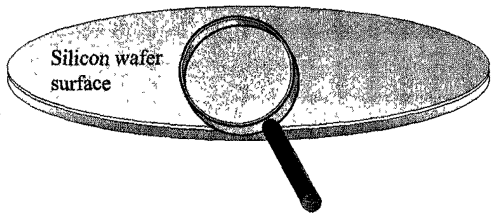


FIGURE 4.28 Flatness of Wafer Front Surface

Microroughness

Microroughness is the small-scale deviation of the actual surface from a nominal plane surface, with many small, closely spaced peaks and valleys.[33] It is an indication of the surface texture on the wafer. Surface microroughness measures the maximum and minimum height deviation on the wafer surface and is specified in units of nanometers (10^{-9} meters). The roughness criteria is usually specified as a root-mean-square (RMS), which is the square root of the average of the square of all deviation measurements from a defined plane. This is a general statistical technique for determining the most probable measurement. Measurement of wafer surface microroughness is done using one of several types of optical surface profile instruments.

Surface microroughness is important to control for wafer fabrication because it is known to have a negative effect leading to the breakdown of very thin dielectric films that are placed on the wafer during device fabrication.[34] Wafers undergo a lapping step followed by an etching step to remove surface microroughness. Typical specification values for microroughness are shown in Table 4.5 on page 85.

Oxygen Content

Control of the oxygen level and uniformity in the silicon ingot is important and becoming increasingly difficult with larger crystal diameters. Small amounts of oxygen can have beneficial gettering effects, which can tie up contaminants in the wafer, as noted previously in this chapter. However, excessive oxygen in the silicon ingot can affect the mechanical and electrical properties of the silicon. For instance, oxygen can lead to an increase in the leakage current at the pn junction and also to increased MOS leakage.[35]

Oxygen content in silicon is verified by using cross-sectioning, which permits composition analysis of the silicon crystal structure. A representative piece of the silicon is potted in epoxy and then ground and polished flat to expose the grain structure of the solid silicon. A chemical etchant is used to darken and highlight certain elements for identification. Once the sample is prepared, special microscopes such as transmission electron microscopy (TEM) are used to characterize the crystal structure. The use of TEM is discussed in Chapter 7. The oxygen content is currently controlled in silicon wafers within a range of 24 to 33 ppm.[36] Typical specifications are shown in Table 4.5 on page 85.

Crystal Defects

It is necessary to control the silicon in order to minimize the various crystal defects previously discussed in this chapter. As shown in Table 4.5 on page 85, the current requirement is for the number of crystal defects per square centimeter to be less than 1000. Cross-sectioning is the technique used to control bulk microdefects in the crystal.

Particles

The number of surface particles are controlled on a wafer to minimize yield loss during wafer fabrication. The primary means of reducing particles is to minimize the generation of particles during wafer processing, followed by effective cleaning steps to remove particles. A typical wafer-cleanliness specification is to have less than 0.13 particles per square centimeter (1300 particles per square meter) of wafer surface area on a 200-mm wafer. This particle size measured is greater than or equal to 0.08 micron.

Bulk Resistivity

The resistivity of the bulk silicon ingot depends on the density of the dopant material that was added to the silicon melt prior to crystal growth. Recall that the most common dopant materials are either boron to create p-type silicon or phosphorus to make n-type silicon. The result of adding these Group IIIA or VA dopants to silicon is to decrease its resistivity due to increased carrier mobility.

It is important to achieve uniform resistivity throughout the bulk silicon. During the actual crystal growth process, there is a radial temperature gradient that reaches a maximum at the center of the ingot and decreases toward the outer edge. This radial temperature gradient produces variations in the radial doping concentration of the ingot.[37]

The bulk silicon ingot is checked for correct resistivity and uniformity after the ends of the ingot have been removed. The four-point probe measurement tool is used to measure wafer resistivity. This measurement tool is described in Chapter 7.

EPITAXIAL LAYER

In some cases, it is desirable to have wafers with a very pure silicon surface of the same crystal structure as the substrate wafer (monocrystalline), yet retain specific control over the doping type or concentration. This condition is achieved by depositing an *epitaxial layer* on the surface of the silicon (referred to as an *epilayer*). Epitaxial is a combination of two Greek words, *epi,* meaning "upon", and *taxis,* meaning "ordered."

In epitaxial silicon, the base wafer is used as a seed crystal to grow a thin layer of silicon on the wafer. The crystal structure of the new epitaxial layer will duplicate that of the wafer. Since the substrate wafer is monocrystal, the epitaxial layer is monocrystal. Furthermore, the dopant of the epitaxial layer can be n-type or p-type and is independent of the initial wafer's dopant type. For example, it is possible to grow a p-type epilayer on a p-type wafer with a lower concentration of electrically active dopant in the epilayer than what is found on the wafer (see Figure 4.29).

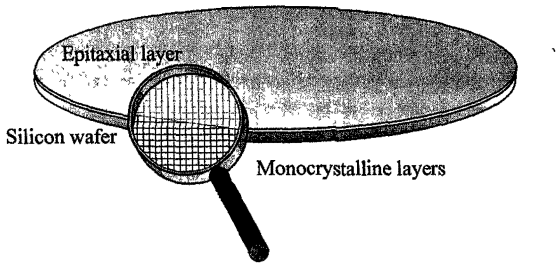


FIGURE 4.29 Formation of Epitaxial Silicon Layers

The original reason for the development of silicon epitaxy was to attain improved performance in bipolar transistors and integrated circuits. Epitaxy permitted the growth of a lightly doped epilayer over a heavily doped substrate, which optimized for a high breakdown voltage in the pn junction while permitting low collector resistance, yielding high device operating speeds at moderate currents. Epitaxy has become important in CMOS ICs because it minimizes latchup problems as device geometries continue to shrink. An epitaxial layer is generally free of contaminants (such as no oxygen particles, which is not true for CZ-grown silicon).

The thickness of the epitaxial layer can vary, with a typical thickness of 0.5 to 5 μm for high speed digital circuit applications, and an epilayer as thick as 50 to 100 μm for silicon power devices.[38] The methods for depositing an epitaxial layer on silicon will be discussed under the topic of deposition in Chapter 11.

SUMMARY

Naturally occurring silicon is used to produce ultrapure, semiconductor-grade polycrystalline silicon. Silicon is a crystal with a repetitive FCC diamond unit cell structure at the atomic level. The crystal orientation is defined by Miller indices, with (100) most common for MOS devices. Polycrystalline silicon is converted into the monocrystal (single-crystal) ingot necessary for wafer fabrication through crystal growth using the CZ method and equipment referred to as a crystal puller. Dopant material is added to the liquid silicon during the CZ process to achieve proper doping levels. Undesirable impurities are tightly controlled in order to grow pure silicon. An alternative process, the float-zone method, produces silicon with very low oxygen content.

Ingot diameters have been increasing over the years to place more chips on a wafer and lower cost through economies of scale. The major crystal defects controlled during growth are point defects, dislocations, and gross defects. Ingots undergo a multistep process to make acceptable wafers, including diameter grinding, wafer notch, wafer slicing, wafer lapping, edge contour, etching, polishing, cleaning, inspection, and packaging.

Wafer suppliers have specific quality measures to control the wafer quality, for characteristics such as physical dimension, flatness, microroughness, oxygen content, crystal defects, particles, and bulk resistivity. An epitaxial layer is sometimes grown on the silicon surface to achieve a very pure silicon layer with the same crystal structure yet have precise control of resistivity by the controlled addition of dopants.

KEY TERMS

semiconductor-grade silicon
crystal
crystal lattice
amorphous materials
unit cell
face-centered cubic (FCC)
FCC diamond structure
polycrystal
polysilicon
monocrystal
single crystal
Miller indices
crystal growth
ingot
Czochralski (CZ) method
CZ crystal puller
melt
magnetic CZ (MCZ)
getter
float-zone method
economies of scale
crystal defect
microdefect
yield defect density
point defect
vacancy (or void)
interstitial
Frenkel defect
substitutional impurities
dislocation
stacking faults
oxygen-induced stacking faults
gross defect
slip
twin planes (twinning)
boule
seed end
tang end

diameter grinding
flats
notch
fixed-quality area (FQA)
wafer slicing
lapping
polished wafer edge (edge grind)
wafer etch
chemical mechanical planarization (CMP)
double-sided polishing (DSP)
wafer flatness
microroughness
epitaxial layer (epilayer)

REVIEW QUESTIONS

1. List the three steps to obtaining semiconductor-grade silicon. How pure is semiconductor-grade silicon?
2. Name the process for producing pure semiconductor-grade silicon?
3. What is a crystal? What is a crystal lattice?
4. Describe an amorphous material. Why is this unacceptable for the silicon used in wafers?
5. Define a unit cell. What type of a unit cell does silicon form?
6. How many complete atoms are in the silicon unit cell? How many atoms are shared and how many are unshared in the silicon unit cell?
7. Describe a polycrystal. What is another name for polycrystalline silicon?
8. Describe a monocrystal. What is another name for a monocrystal?
9. Why is it necessary to have monocrystal silicon for wafer fabrication?
10. What are Miller indices, and what do they indicate?
11. Draw a picture of the following three crystal planes: (100), (110) and (111).
12. Which crystal plane orientation is most common for MOS? Which is most common for bipolar?
13. Define crystal growth.
14. What is the CZ method for crystal growth?
15. Describe the silicon seed for the CZ method and how it is used.
16. What is the name of the main equipment used in the CZ method?
17. Describe a crystal puller. Specifically, describe the crucible holder, seed, melt, pull process and ingot growth.
18. What is the objective of the pull process and what are the two main parameters that affect the pull process in the CZ method?
19. What is the purpose of magnetic CZ?
20. Why is dopant material added to the melt during the CZ method?
21. How can small amounts of oxygen in the ingot be beneficial?
22. Why is the float-zone method desirable to grow silicon crystal?
23. Describe the float-zone process.
24. What is the main reason for the continued increase in wafer diameter?
25. Give three benefits to using larger-diameter wafers.
26. What is a crystal defect?
27. Define defect density. How is defect density related to wafer yield?
28. What are the three general forms of crystal defects in silicon?
29. List and describe three types of point defects.
30. Which crystal growth parameters can affect the occurence of point defects.
31. What is a dislocation? Explain a stacking fault.
32. What is an oxidation-induced stacking fault?
33. Explain how crystal growth dislocations have been reduced for large-diameter wafers.
34. What is a gross defect? Explain the difference between crystal slip and crystal twin planes.
35. What is the purpose of diameter grinding of the ingot?
36. Describe or draw a picture of the four types of wafer flats. What have the wafer flats been replaced with on 200-mm diameter and larger wafers?
37. How is wafer slicing done on wafers up to 200 mm in diameter?
38. What is the purpose of wafer lapping?
39. Why is a polished wafer edge (edge grind) done to the wafer? Give three reasons why this grinding benefits the wafer quality.
40. Why is chemical etching done on the wafer?
41. Why is chemical mechanical planarization (CMP) done on the wafer surface?
42. How are wafers packaged for shipment to the wafer fab?
43. List seven wafer quality requirements for a silicon wafer.
44. What is wafer flatness, and how is it measured?
45. What is surface microroughness, and why is it important to control?
46. What is an epitaxial layer, and why is it used on wafers?

SELECTED INDUSTRY WEB SITES

GT Equipment Technologies, Inc.	http://www.gtequipment.com/
International SEMATECH	http://www.sematech.org/public/index.htm
Kayex CZ Crystal Growers	http://www.kayex.com/
Mitsubishi	http://www.munc-sil.com/
NIST, National Institute of Standards and Technology	http://www.nist.gov/
Selete, Semiconductor Leading Edge Technologies, Inc.	http://www.selete.co.jp
SEMI, Semiconductor Equipment and Materials International	http://www.semi.org/
Semiconductor International Magazine	http://semiconductor.net/
Semiconductor Search Engine	http://www.semiseek.com/
SIA, Semiconduictor Industry Association	http://www.semichips.org/
SISA, Semiconductor Industry Suppliers Association	http://www.sisa.org/
Solid State Technology Magazine	http://sst.pennwellnet.com/home/home.cfm
Wafer World, Inc.	http://www.waferworld.com/

REFERENCES

1. C. Pearce, *Crystal Growth and Wafer Preparation,* VLSI Technology, 2nd ed., ed. S. Sze (Boston: McGraw-Hill, 1988), p. 11.
2. S. Ghandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide,* 2nd ed., (New York: Wiley, 1994), p. 102.
3. K. Bachmann, *The Materials Science of Microelectronics* (New York: VCH Publishers, 1995), p. 102.
4. Society of Chemical Engineers of Japan, eds., *Introduction to VLSI Process Engineering* (New York: Chapman and Hall, 1993), p. 148.
5. L. Van Vlack, *Elements of Materials Science and Engineering,* 4th ed., (Reading: Addison-Wesley, 1980), p. 73.
6. S. Ghandhi, *VLSI Fabrication Principles,* p. 16.
7. S. Wolf and R. Tauber, *Silicon Processing for the VLSI Era,* vol. 1 of *Process Technology* (Sunset Beach: Lattice Press, 1986), p. 2.
8. S. Ghandhi, *VLSI Fabrication Principles,* p. 21.
9. Ibid.
10. S. Takasu and W. Zulehner, "Silicon Crystal Growth," *Semiconductor Silicon Proceedings* 94–10 (Pennington: The Electrochemical Society, 1994): p. 55.
11. K.-M. Kim, "Bigger and Better CZ Silicon Crystals," *Solid State Techonology* (November 1996), p. 73.
12. K. Bachmann, *Materials Science of Microelectronics,* p. 248.
13. S. Sze, *VLSI Technology,* 2nd ed. (New York: McGraw-Hill, 1988), p. 26.
14. K. Bachmann, *Materials Science of Microelectronics,* p. 270.
15. S. Ghandhi, *VLSI Fabrication Principles,* p. 120.
16. K. Bachmann, *Materials Science of Microelectronics,* p. 248.
17. K.-M. Kim, "Bigger and Better CZ Silicon Crystals," p. 74.
18. Ibid., p. 74.
19. H. Wolfgang, "High Purity Silicon IV," *International Symposium,* (Pennington, NJ: The Electrochemical Society, 1996): p. 5.
20. H. Huff et al., "Thermal Processing Issues for 300 mm Silicon Wafers: Challenges and Opportunities," *ULSI Science and Technology: Proceedings of the Electrochemical Society,* 97-3: p. 136.
21. Ibid., p. 136.
22. K.-M. Kim, "Bigger and Better CZ Silicon Crystals," p. 74.
23. B. El-Kareh, *Fundamentals of Semiconductor Processing Technologies,* (Boston: Kluwer Academic Publishers, 1995), p. 68.
24. A. Takao, "Innovative Silicon Crystal Growth and Wafering Techonologies," *ULSI Science and Technology: Proceedings of the Electrochemical Society,* 97-3: p. 123.
25. S. Brunkhorst and D. Sloat, "The Impact of the 300-mm Transition on Silicon Wafer Suppliers," *Solid State Technology* (January 1998): p. 90.
26. Ibid.
27. G. Fisher, "Challenge for 300mm Polished Wafer Manufacturers," *Semiconductor International* (September 1998): p. 98.
28. A. Takao, *Innovative Silicon Crystal Growth,* p. 124.
29. H. Huff, et al., "Thermal Processing Issues," p. 140.
30. S. Sze, *VLSI Technology,* p. 40.
31. S. Brunkhorst and D. Sloat, "Impact of the 300-mm Transition," p. 88.
32. G. Fisher, "Challenges," p. 102.
33. W. Runyan and T. Shaffner, *Semiconductor Measurements and Instrumentation,* 2nd ed., (New York: McGraw-Hill, 1997), p. 214.
34. W. Bullis, "Microroughness of Silicon Wafers," *Proceedings of Semiconductor Silicon* 94-10 (Pennington, NJ: The Electrochemical Society, 1994): p. 1156.
35. J. Vanhellemont et al., "On the Electrical Activity of Oxygen-Related Extended Defects in Silicon," *Semiconductor Silicon Proceedings* 94–10 (Pennington, NJ) The Electrochemical Society, 1994): p. 670.
36. K.-M. Kim, "Bigger and Better CZ Silicon Crystals," p. 73.
37. S. Ghandhi, *VLSI Fabrication,* p. 258.
38. S. Ghandhi, *VLSI Fabrication,* p. 258.

CHAPTER 5

CHEMICALS IN SEMICONDUCTOR FABRICATION

The fabrication of semiconductors is largely a chemical process. Chemical manufacturers obtain raw chemicals from nature, refine the chemical makeup of these materials to achieve ultrahigh purity, and transport them to the fabrication area, where these chemicals are introduced to the surface of the wafer. Chemical reactions alter the silicon to create semiconductors and then apply many layers of interconnections. These chemical reactions form the basis for semiconductor device fabrication and microchip performance.

The semiconductor industry is undergoing rapid changes to introduce fundamentally new materials and processes into wafer fabrication. The goal is to achieve improved microchip performance and productivity. Examples of new technologies are insulators with a low dielectric constant (see Chapter 11) and copper metallurgy (see Chapter 12). Chip design and process improvements create the ongoing need for new and improved chemicals.

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Identify and discuss the four states of matter.
2. Describe the important chemical properties relevant to semiconductor manufacturing.
3. State how the different process chemicals are categorized and used in a wafer fab.
4. Explain how an acid, a base, and a solvent are used in chip manufacturing.
5. State whether a gas is a bulk or specialty gas and explain how each type of gas is delivered and used in wafer fabrication.

INTRODUCTION

Semiconductor manufacturing makes use of large quantities of chemicals to fabricate wafers. There is also a high usage of chemicals for cleaning wafers and the processing tools used in the manufacturing processes. Chemicals used in wafer fabrication are referred to as *process chemicals.* They exist in various chemical states and are controlled to strict purity requirements. To understand process chemicals, we will first review some concepts in basic chemistry.

STATES OF MATTER

All matter in the universe exists in three basic states: solid, liquid, or gas. There is an additional fourth state that is not widely understood, known as plasma. We can describe each state by the way the substance fills a container (see Figure 5.1 on page 92).

A *solid* has its own fixed shape and will not conform to the shape of the container. A *liquid* conforms to the container shape. It will fill the container to the extent of the liquid's volume, and has a surface. A *gas* also conforms to the container shape, but it fills the entire container volume

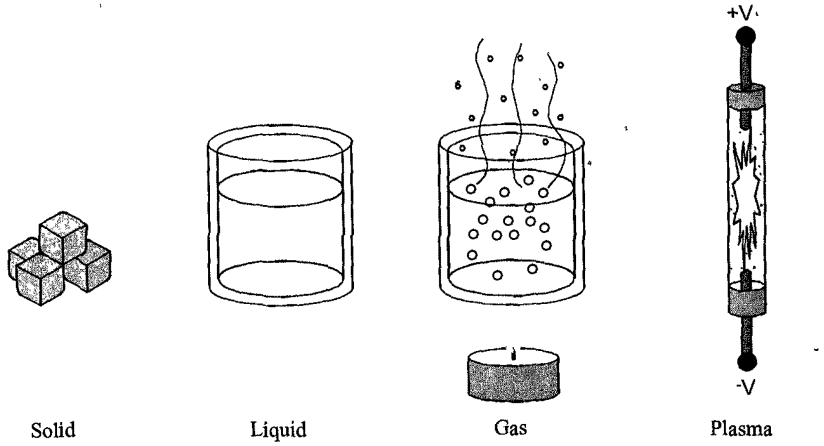


FIGURE 5.1 Physical States of Matter

without forming a surface. Gas particles are small and freely moving. Some gases, like hydrogen and oxygen, are reactive and easily form stable chemical compounds with other gases or elements. Other gases, like helium and argon, are inert. An *inert gas* (also called a *noble gas*) does not readily form compounds. Inert gases are widely used in semiconductor manufacturing because they do not react with other chemicals.

A fourth state of matter is *plasma.* A plasma state exists when there is a high-energy collection of ionized molecules or atoms. This state is considerably different than the three basic states of matter. Examples of a plasma state are a star, a fluorescent light, and neon signs. Plasma states can be induced in certain gases through exposure to high-energy electric fields. As we shall see in later chapters, plasma states are widely used in semiconductor manufacturing.

Many substances can exist in any one of the three basic physical states. For example, water undergoes changes of state depending on the temperature and pressure of its surroundings. As temperature increases, solid water (ice) melts to a liquid water, and then boils into a gas state (water vapor). Many substances behave in this same manner.

PROPERTIES OF MATERIALS

We can understand how to use materials to our benefit in semiconductor manufacturing by investigating their properties. *Properties* are the characteristics of a material that describe its unique identity. For instance, a solid material can have different properties, such as a rigid solid (iron), a soft solid (wax) at room temperature (25°C, 77°F) and a flexible solid (lead). Identifying properties for materials used in semiconductor manufacturing is important for understanding how to properly fabricate the silicon wafer to build chips.

There are two types of material properties: physical and chemical. *Physical properties* of a material are those that reflect a material by itself without its interacting with another substance. Some of these physical properties are the melting point, boiling point, resistivity, and density. *Chemical properties* of a material result from an interaction or transformation with another substance. Examples of chemical properties are flammability, reactivity, and corrosiveness. A *chemical reaction* occurs when a material is converted into a different material that has a different composition and properties. The starting material is called a *reactant,* and the resultant material is called a *product.*

An example of a chemical reaction is the combustion of hydrogen in the presence of oxygen. Both chemicals exist as gases in their natural state. Hydrogen will undergo a chemical reaction with oxygen when the temperature of hydrogen exceeds 600°C. The result is an explosive heat reaction with water vapor as the byproduct. The chemical change is:

$$2 \text{ parts of hydrogen (gas)} + \text{oxygen (gas)} \xrightarrow{\text{Heat}} \text{water (liquid)} + \text{energy}$$

Chemical Properties for Semiconductor Manufacturing

There are many different types of chemicals and materials used to fabricate silicon microchips. Manufacturers of advanced ICs use new materials to improve chip performance and reduce device feature sizes. Some chemical properties are important to understanding how existing and new semiconductor process materials perform. These properties are:

- Temperature
- Pressure and vacuum
- Condensation
- Vapor pressure
- Sublimation and deposition
- Density
- Surface tension
- Thermal expansion
- Stress

Temperature ■ *Temperature* is a measure of how hot or cold a substance is relative to another substance. Thus it is a measure of the average kinetic or thermal energy (molecular or atomic motion) of a substance. The transfer of energy between objects at different temperatures is *heat.* Wafer fabrication requires extensive processing at high temperatures for reasons such as using heat to affect chemical reactions (e.g., changing the rate of a reaction) or annealing the silicon crystal structure to rearrange atoms.

Three temperature scales exist: Fahrenheit in °F, Celsius in °C (commonly called centigrade), and Kelvin in K. Figure 5.2 illustrates how these three scales are related.

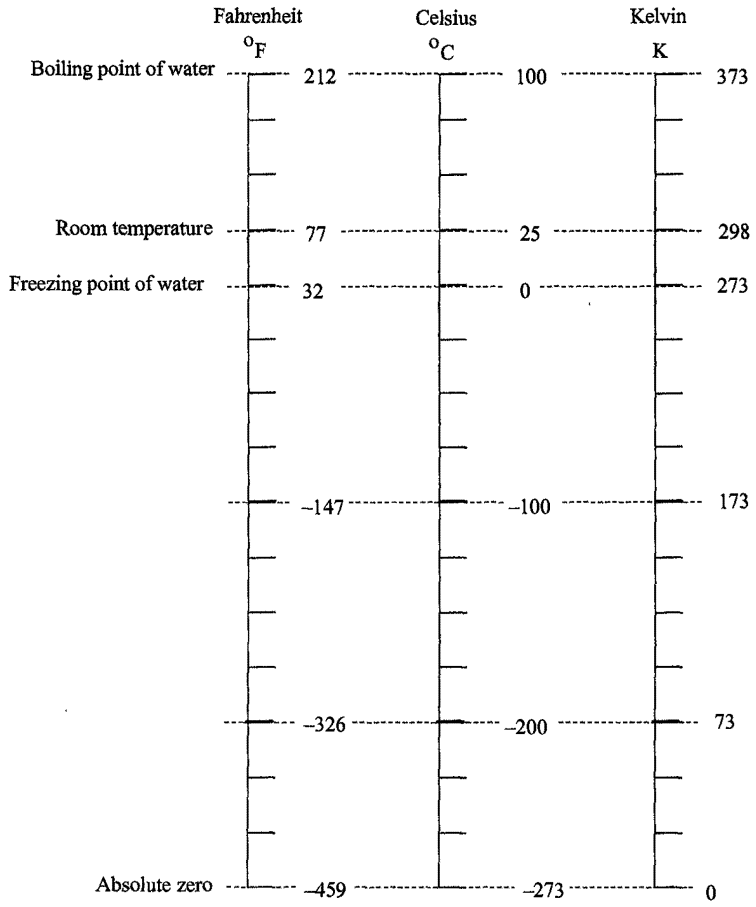


FIGURE 5.2 Temperature Scales

The most common temperature scale in scientific work is the *Celsius* scale. This scale is based on changes in the physical state of water: 0°C is set at the water's freezing point and 100°C at its boiling point when measured at standard atmospheric pressure at sea level. To convert between Fahrenheit and Celsius,

$$°F = \tfrac{9}{5}\,°C + 32$$
$$°C = \tfrac{5}{9}[°F - 32]$$

The *Kelvin* scale, or absolute scale, is the base unit of temperature for the metric unit system. The temperature 0 K (absolute zero) is −273°C (more precisely, −273.15°C), which is the lowest temperature attainable and the theoretical temperature at which all atomic motion would cease. To convert between Kelvin and Celsius,

$$K = °C + 273$$
$$°C = K - 273$$

Pressure and Vacuum of a Gas ■ Gas fills the entire volume of a container and exerts uniform pressure on its walls. *Pressure* (P) is defined as the force exerted per unit area against a surface:

$$\text{Pressure} = \frac{\text{Force}}{\text{Area}} \text{ (pounds per square inch, or psi)}$$

The pressure of any gas is dependent on the number of gas molecules present, the temperature, and the volume of the chamber. If there are more gas molecules striking the sides of the container per unit time, then pressure increases (see Figure 5.3). This is also evident when blowing up a balloon.

The English unit of pressure common in the United States is pounds per square inch (psi), specified as either psia (absolute pressure, which includes atmospheric pressure of 14.7 psi) or psig (gauge pressure). These two gauges are illustrated in Figure 5.4. Different units of pressure at standard atmosphere and temperature (sea level and 23°C) are shown in Table 5.1.

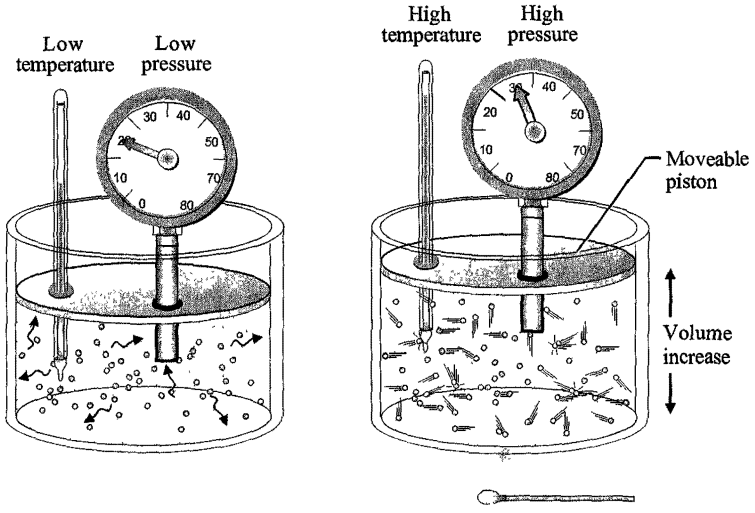


FIGURE 5.3 Pressure Against a Container Wall

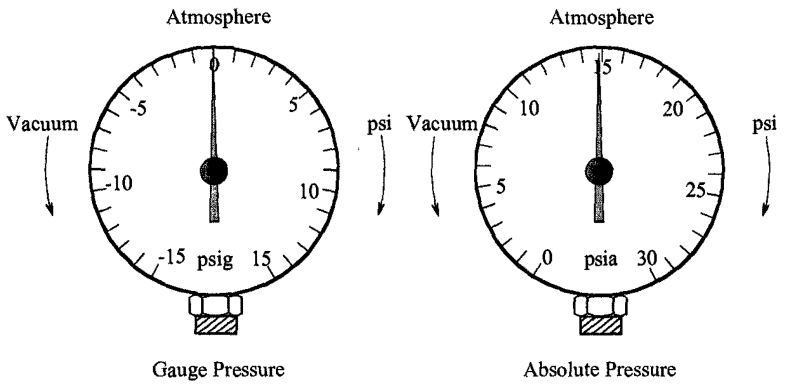


FIGURE 5.4 Gauge Pressure (PSIG) Versus Absolute Pressure (PSIA)

TABLE 5.1 Units of Pressure at Sea Level and 23°C

Unit of Pressure	Sea Level and 23°C
Psig (gauge)	0 psi
Psia (absolute)	14.7 psi
Atmosphere	14.7 psi
Inches of mercury	29.92 inches
Millimeters of mercury	760 mm
Torr	760 torr
mtorr	760,000 mtorr
Bar	1.013 bar
millibar	1013 mbar
Pascal	101,325 pascal

Pressure is a widely used property in semiconductor manufacturing. Chemicals and gases flow from regions of higher pressure to lower pressure. Some fabrication processes occur at atmospheric pressure, while some require higher pressures. Other processes occur in chambers at less than atmospheric pressure (vacuum).

Vacuum. If the gas pressure in a container is less than 14.7 psi, then a vacuum exists. *Vacuum* is the removal of gas molecules (e.g., air, moisture, and gas residues) in a closed container to achieve a pressure less than atmosphere. Many semiconductor operations occur in vacuum conditions. Vacuum is commonly measured in units of torr. A *torr* is defined as the equivalent of one mm of mercury in a pressure-measuring device known as a barometer (see Figure 5.5 on page 96).

At one atmosphere, the weight of the air will push down on the mercury in the bowl and cause it to rise 760 mm (760 torr or 29.92 inches) in the vacuum of the column. The amount of mercury rise is proportional to the pressure bearing down on the mercury in the dish. Less pressure (or vacuum) makes the mercury rise less. For the sensitivity required to measure and control vacuum conditions in many semiconductor processes, the units of millitorr are required.

Condensation and Vaporization ■ The process of a gas changing into a liquid is called *condensation*. When water vapor cools, a mist appears as the particles form tiny droplets of liquid that then collect into a bulk sample with a single surface. The opposite process, changing from a liquid into a gas, is called *vaporization.*

Liquids and gases interact with materials in different ways. *Absorption* is the taking up of a liquid or a gas into the bulk of another material, as in the dissolving of a gas in a liquid. *Adsorption* is the condensing of a gas or a liquid on the surface of a solid. The adsorbed molecules actually stick to the surface through a chemical bond or through a weaker bond of physical attraction.

Vapor Pressure ■ *Vapor pressure* is the pressure exerted by vapor in a closed container when there is an equilibrium condition between the rate of vaporization and rate of condensation. Vapor pressure is shown in Figure 5.6 on page 96.

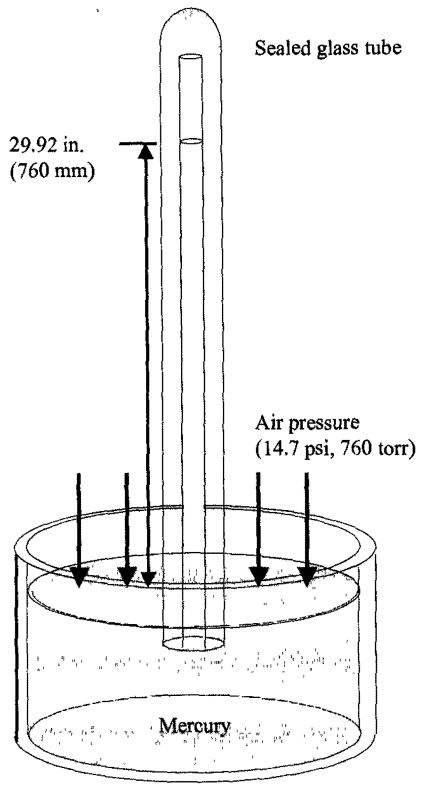


FIGURE 5.5 Barometer at Atmospheric Pressure

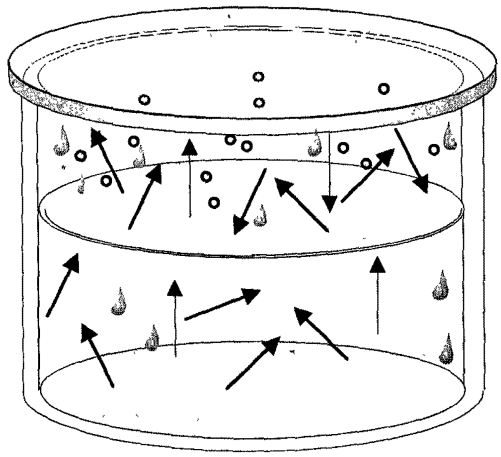


FIGURE 5.6 Vapor Pressure

High vapor pressure materials are *volatile* (tend to become a gas). A material with a high vapor pressure will release vapors more easily when exposed to vacuum conditions. Examples are solvents, perfumes, and lotions. These materials typically exhibit odors that are readily detectable through human nasal passages.

Sublimation and Deposition ■ A solid can change directly to a gas through a process termed *sublimation*. Have you ever put an ice tray in the freezer and left it there for six months? The ice cubes are smaller because the ice sublimes. Other examples of sublimation are dry ice (shown in Figure 5.7) and moth balls.

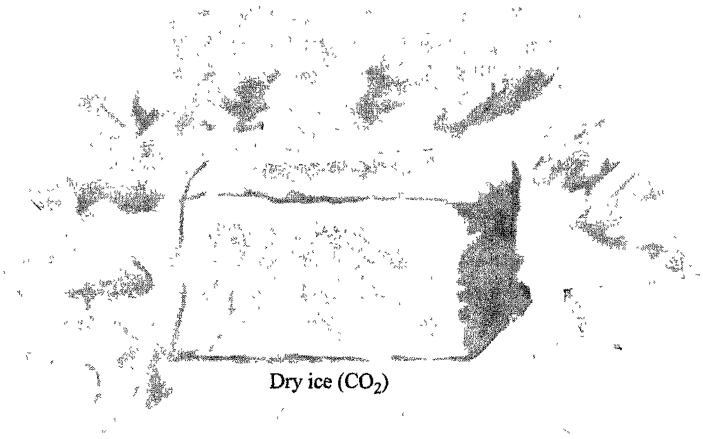


FIGURE 5.7 Sublimation

The opposite process, changing from a gas into a solid, is called *deposition* (see Figure 5.8). This process is how ice forms on a cold window from the deposition of water vapor. As we shall see in a later chapter, deposition is an important process in semiconductor manufacturing.

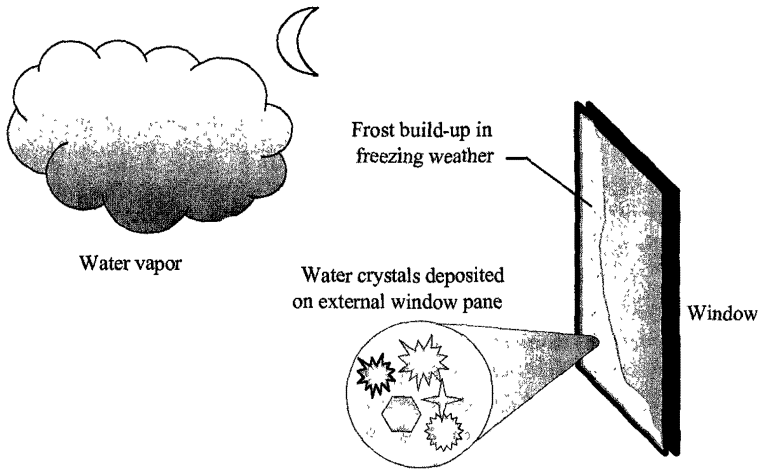


FIGURE 5.8 Deposition

Density ■ The *density* of a substance is defined as its mass (or weight) divided by its volume:

$$\text{Density} = \frac{\text{Mass}}{\text{Volume}} \text{ (grams/cm}^3\text{)}$$

A dense material is thought of as being heavy. If two objects have the same volume and one object is heavier than the other is, then it is denser (see Figure 5.9). A sponge has a lower density than an equal volume of steel. We know this because the sponge weighs less than the steel.

Table 5.2 on page 98 gives the density for different substances at standard temperature and pressure.[1] Water is the standard, with a density of 1 gram/cm^3. The densities of other substances are often expressed as a ratio of their density to that of a comparable volume of water. For instance, silicon has a density of 2.3. Interpret this number to mean that a piece of silicon occupying a volume of one cm^3 will weigh 2.3 grams.

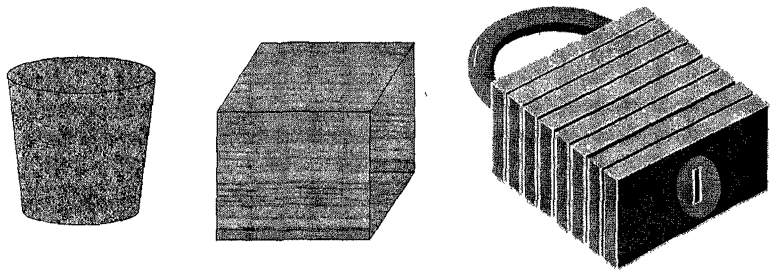


FIGURE 5.9 Density of Objects

TABLE 5.2 Densities of Some Common Substances at Standard Temperature and Pressure

Substance	Physical State	Density (g/cm^3)
Hydrogen	Gas	0.000089
Oxygen	Gas	0.0014
Water	Liquid	1.0
Table salt	Solid	2.16
Silicon	Solid	2.33
Aluminum	Solid	2.70
Gold	Solid	19.3

A term related to density is specific gravity. *Specific gravity* (*SG*) refers to the density of liquids and gases at 4°C and is the ratio of the density of a substance to the density of water.

Surface Tension ■ When a liquid is on a flat surface, it has a surface area of contact (see Figure 5.10). The *surface tension* of the liquid is the energy required to increase the surface area of contact. For the surface area to increase, liquid molecules must break intermolecular attraction forces and move away from the interior of the liquid and toward its surface. This movement requires energy. The concept of surface tension is used in semiconductor manufacturing to measure the ability of liquid coatings to uniformly adhere to the wafer surface.



FIGURE 5.10 Surface Tension of a Liquid on a Wafer

Thermal Expansion ■ When an object is heated, it expands due to the increased vibrations of the atoms. Due to this thermal expansion, the dimensions of a heated object will actually increase (see Figure 5.11). When the same object is cooled, its dimensions will decrease.

Some materials will expand thermally more than others. The amount a material expands due to heating is known as its *coefficient of thermal expansion,* or *CTE.* Amorphous materials that have a cubic crystal orientation expand to thermal exposure in all directions. For all other crystals, such as monocrystal silicon, thermal expansion varies with the type of crystal orientation.

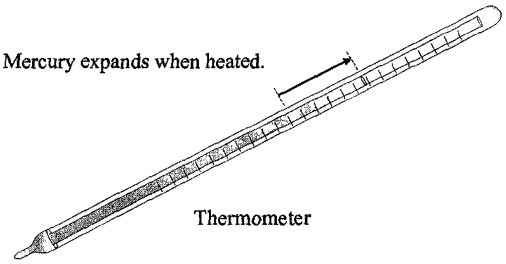


FIGURE 5.11 Thermal Expansion of a Heated Object

Stress ■ *Stress* occurs when an object is exposed to a force. The magnitude of the stress depends on two factors: the amount of force, and the area over which the force is applied. The units of stress are pounds per square inch (psi) or pascal (Pa) in the metric system. The formula for stress is,

$$\text{Stress} = \frac{\text{Force}}{\text{Area}} \text{ (psi)}$$

The force that creates stress in a wafer can come from various sources. Stress can come from work damage on the surface of the wafer; from internal forces due to dislocations, excess vacancies and impurities; and from growth around included foreign material.[2] If two dissimilar materials with significantly different coefficients of thermal expansion (CTE) are bonded together and heated, then

a force occurs because the two materials expand at different rates and pull against each other to cause a stress. This stress can create a warped wafer due to the CTE mismatch. Such stress is a concern in semiconductor manufacturing because a microchip is a planar structure with layers of different materials with different CTEs. Deposited films usually have some form of stress that can be compressive or tensile, and the nature of the stress depends on process conditions (see Figure 5.12). The reliability of a chip is improved by ensuring materials have minimal stress.

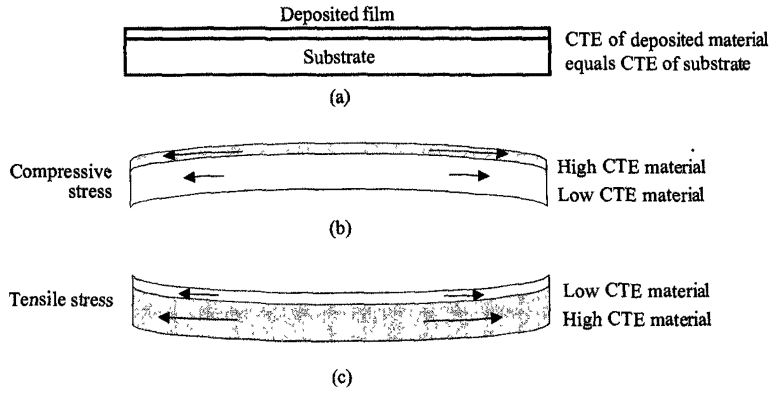


FIGURE 5.12 CTE Mismatch of Two Materials

PROCESS CHEMICALS

Semiconductor manufacturing is a chemically intensive process involving many different types of process chemicals of ultrahigh purity. The process chemicals are used in all three states: liquids, solids, and gases. Chemicals are used in the process of manufacturing semiconductors to:

- Clean or prepare the wafer surface with wet chemical solutions and ultrapure water rinse.
- Dope the wafer with energetic atoms to create p-type and n-type silicon.
- Deposit the different metal conductor layers and the necessary dielectric layers between the conductors.
- Grow the thin silicon dioxide film to be used as the critical MOS gate dielectric.
- Etch thin films with plasma or wet chemicals to selectively remove material and form the required pattern in the film.

There is a vast array of chemicals used in the fab, yet some predominant chemicals are repeatedly used in the different process areas of the fab. Cleaning is one of the most chemically intense steps of semiconductor manufacturing. It is estimated that 30% of all processing steps in the wafer fabrication process are for wafer cleaning or wafer surface preparation.[3]

Liquids

Liquids can be either a pure substance, such as pure water, or a chemical mixture. Automotive gasoline is an example of a chemical mixture of hydrocarbons and additives blended together to produce efficient combustion.

If a chemical mixture is so well mixed that the molecules or ions of the components are the same throughout, then we call this a *chemical solution.* Gasoline is a chemical solution. Another example of a solution is the household first aid antiseptic hydrogen peroxide, which consists mostly of water, with only 5% hydrogen peroxide. The component of the solution present in the larger amount (such as water in the hydrogen peroxide solution) is called the *solvent.* The dissolved substances in the solution are the *solutes.* Solutions in water are called *aqueous solutions,* meaning water is the solvent.

There are many types of liquids used in the wet processes of semiconductor manufacturing. All liquids used in wafer fabs must have extremely high purity, with no contamination from particulates, metallic ions, or unwanted chemicals. Contamination of chemicals is a relative

concept. A term that is frequently used to describe small concentrations of an impurity is *parts per million* (*ppm*), which is measured by volume or by mass (weight). An example of ppm is if you consider the amount of an impurity in air. To calculate the ppm of an impurity by mass, you determine the weight of the impurity in a fixed volume of air, divide this number by the weight of the air, and multiply by one million. Chemical purity specifications have become more stringent so that ultrapure chemicals used in the fab often have less than a *parts per billion* (*ppb*) or *parts per trillion* (*ppt*) of an impurity.[4] The amount of impurity measured in a wafer fab process chemical is often limited by the measurement capability of the instrument.

There has been a long-term effort to reduce the quantity of liquid chemicals in the wafer fab. Many liquid chemicals used in semiconductor manufacturing are hazardous and require special handling and disposal procedures. Furthermore, chemical residues can contaminate wafers, as well as produce vapors that diffuse through the air and deposit on wafer surfaces. Although it has not been possible to eliminate wet chemicals from the wafer fab, there has been a significant reduction in their use, such as in reducing the number of wet cleaning steps.[5]

Liquid process chemicals in a wafer fab are grouped in the following manner:

- Acid
- Base
- Solvent

Acid ■ In the classical definition, *acids* are solutions that contain hydrogen and dissociate in water (meaning bonds break down) to yield hydronium ions, H_3O^+. Since an acid contains hydrogen, it contains H in its formula, such as phosphoric acid (H_3PO_4) or hydrochloric acid (HCl).

As an example of how acid dissociates in water, consider the following reaction for hydrochloric acid:

$$HCl\ (gas) + H_2O\ (liquid) \rightarrow Cl^-\ (aqueous) + H_3O^+\ (aqueous)$$

We can see that when HCl is dissolved in water, it reacts with the water to produce the hydronium ion, H_3O^+. This defines HCl as an acid.

There are many different acids used in semiconductor manufacturing. Table 5.3 lists some common acids and their typical use in a wafer fab.

TABLE 5.3 Common Acids Used in Semiconductor Manufacturing

Acid	Symbol	Example of Use*
Hydrofluoric acid	HF	Etching silicon dioxide (SiO_2) and cleaning quartzware
Hydrochloric acid	HCl	Wet cleaning chemical that is part of the standard clean 2 (SC-2) solution to remove heavy metals from wafer
Sulfuric acid	H_2SO_4	Solution known as "Piranha" [7 parts H_2SO_4 to 3 parts of 30% hydrogen peroxide (H_2O_2)] used to clean wafers
Buffered oxide etch (BOE): Solution of hydrofluoric acid and ammonium fluoride	HF and NH_4F	Etching of silicon dioxide (SiO_2) film
Phosphoric acid	H_3PO_4	Etching of silicon nitride (Si_3N_4)
Nitric acid		HNO_3Used in mixture of HF and HNO_3 to etch phosphosilicate glass (PSG)

*Explanation of wafer cleaning is in Chapter 6 and etching in Chapter 16.

Acids are further divided into two categories: organic and inorganic. Organic acids such as carboxylic acids contain hydrocarbons, while inorganic acids such as hydrofluoric acid (HF) do not.

Base ■ A *base* is a substance that contains the OH chemical group (e.g., NaOH, sodium hydroxide and KOH, potassium hydroxide) and dissociates in water to yield the hydroxide ion, OH^-. Another name for a base substance is *alkali* or an *alkaline substance*. A base will increase the

hydroxide ion in an aqueous solution. For example, sodium hydroxide is an ionic compound containing metal ions and hydroxide ions. It is a base because when it dissolves in water it dissociates to yield Na^+ and OH^- ions:

$$NaOH \rightarrow Na^+ + OH^-$$

Common alkaline substances used in semiconductor manufacturing are shown in Table 5.4.

TABLE 5.4 Common Bases Used in Semiconductor Manufacturing

Base	Symbol	Example of Use
Sodium hydroxide	NaOH	Wet etchant
Ammonium hydroxide	NH_4OH	Cleaning solution
Potassium hydroxide	KOH	Positive photoresist developer
Tetramethyl ammonium hydroxide	TMAH	Positive photoresist developer

pH ■ An acid or a base varies in strength and is classified as strong or weak. The *pH scale* is used to assess how strong or weak a solution is as an acid or base. This scale ranges from 0 to 14, with 7 being the neutral point. Acids have pH values below 7 and bases have pH values above 7. Pure water is the reference substance for the pH scale and is neutral with a pH of 7. Strong acids, such as sulfuric acid (H_2SO_4) have low pH values between 0 and 3. Strong bases, such as sodium hydroxide (NaOH) have pH values greater than 7 and approaching 14. Figure 5.13 shows some common chemicals and their location on the pH scale.

pH	Household Chemicals	
1	Car battery acid (sulfuric acid, H_2SO_4)	Corrosive
2		
3	Lemon juice, vinegar	
4	Soda, wine	
5	Tomato juice, beer	
6	Urine	
7	Tap water, milk, saliva	
8	Blood, saliva	
9	Milk of magnesia	
10	Detergents	
11	Household ammonia	
12		
13	Household drain cleaners	Caustic
14	Nickle-cadmium battery (NaOH base)	

FIGURE 5.13 The pH Scale for Different Chemicals

Solvents ■ A *solvent* is a substance capable of dissolving another substance to form a solution. A good solvent will dissolve (solubilize) a broad range of substances. Most solvents, such as alcohol and acetone, are volatile and flammable. Common solvents used in a wafer fab are listed in Table 5.5 on page 102.

TABLE 5.5 Common Solvents Used in Semiconductor Manufacturing

Solvent	Common Name	Example of Use
Deionized water	DI Water	Widely used to rinse wafers and to dilute cleaning solutions.
Isopropyl alcohol	IPA	General-purpose cleaning solvent.
Trichloroethylene	TCE	Solvent used for wafer and general cleaning.
Acetone	Acetone	General-purpose cleaning solvent (stronger than IPA).
Xylene	Xylene	Strong cleaning solvent; may also be used for photoresist edge bead removal.

Deionized water (DI water) is a widely used solvent in semiconductor manufacturing that has all conductive ions removed. With a pH of 7, which is neutral, DI water is neither an acid nor a base. It has the ability to dissolve other substances, including many ionic and covalent compounds. When H_2O water molecules dissolve an ionic compound, they separate, surround, and disperse the ions into the liquid. Water does this by overcoming the electrostatic force of attraction between the ions.

Chemical Distribution ■ There is a wide range of chemicals used in the semiconductor industry, with many being toxic and hazardous. A review of key chemical safety issues is provided in Appendix 1. Safe, high-purity, uninterrupted delivery of chemicals from storage vessel to process tools is critical. For liquid chemicals, this delivery is often accomplished through bulk chemical distribution (BCD).

A *bulk chemical distribution* system consists of a chemical source, such as a storage vessel, a chemical delivery module, and a piping system (see Figure 5.14).[6] The BCD storage vessel is often housed below the main production floor in the groundlevel sub-fab. The chemical delivery module will filter, blend, and transport the chemicals through the piping system. A piping system delivers the chemical to the individual process stations. Modern BCDs are integrated into a fully computerized and networked system for real-time chemical monitoring and control.

The decision regarding how to store and deliver process chemicals depends on factors such as chemical compatibility, reduction of chemical contamination, and safety. Purity requirements of

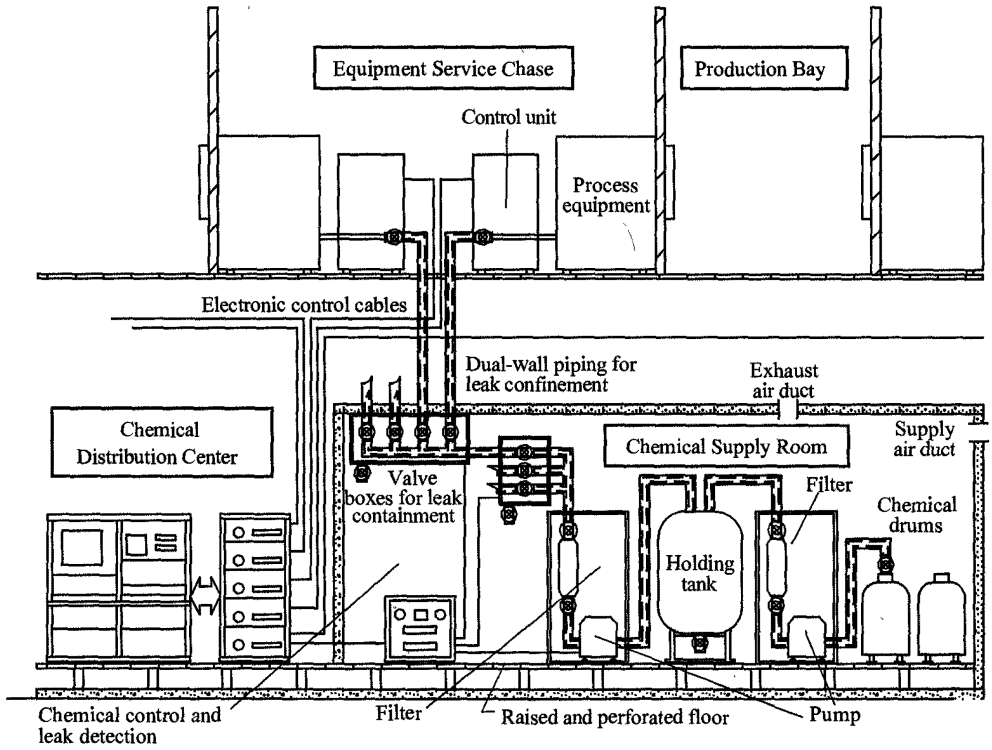
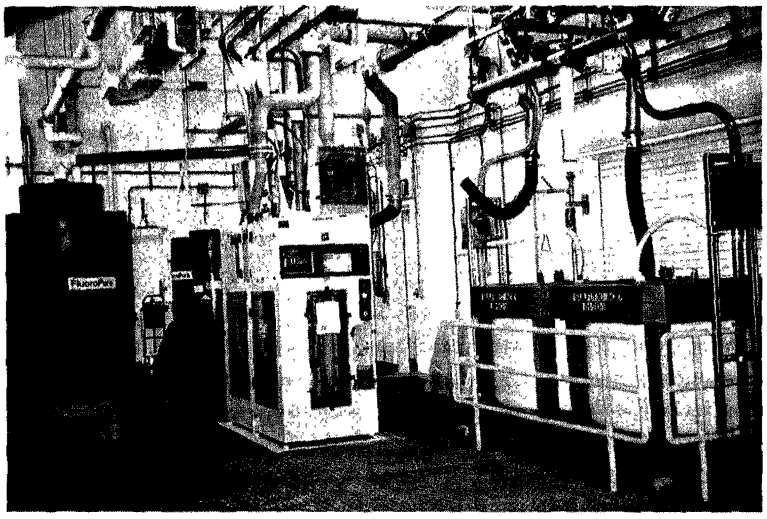


FIGURE 5.14 Bulk Chemical Distribution



Bulk Chemical Distribution
(Photo courtesy of Advanced Micro Devices)

chemicals used in semiconductor manufacturing are described as *ultrahigh purity* (*UHP*), with contaminants being controlled to the parts per billion (ppb) to parts per trillion (ppt) range.

Some chemicals are not suitable for BCD. They may be used in small quantities or have a limited shelf life (how long they can be stored before use). In this case, these chemicals will have special packaging systems for *point-of-use* (*POU*) delivery, which means they are stored and used at the process station. An example of this type of chemical is the photoresist chemical used in photolithography (discussed in Chapter 13).

Gases

There are about 50 different types of gases used throughout the 450 or so process steps in semiconductor manufacturing.[7] The types and quantities of gases are changing because there are new materials being introduced into the fab process, including copper metallurgy and new diffusion barriers (see Chapter 12). Gases are usually categorized as one of two types: bulk gases or specialty gases. *Bulk gases* are traditionally defined as oxygen (O_2), nitrogen (N_2), hydrogen (H_2), helium (He), and argon (Ar). The *specialty gases* are also referred to as process gases, and are the other important gases used to manufacture semiconductor microchips.

Extremely high purity is required of all gases, with bulk gases controlled to seven nines purity (99.99999%) and specialty gases controlled at least to four nines (99.99%). Particulate contamination in gases is controlled to the 0.1-μm range. Other controlled contaminants are oxygen, moisture content, and trace impurities such as metals.[8] Many process gases are toxic, corrosive, reactive, and pyrophoric (they combust or burn when exposed to air). For these reasons, gases are contained in a gas delivery system that delivers the gases in a safe, clean, and accurate manner to the various process stations in the wafer fab.

Bulk Gases ■ Bulk gases are relatively simple gases for the gas suppliers to manufacture and are stored outside of the wafer fab manufacturing area in large storage tanks or 1000-lb bulk tube trailers. These gases are distributed to the workstation through the bulk gas distribution (BGD) system. Benefits from the centralized gas control of BGD are a reliable gas supply, less sources of particulate contamination, and less human involvement in the daily delivery of the gas. Bulk gas is often the lowest cost method for gas delivery with higher gas purity. Furthermore, on-site production of the bulk gases is done at large wafer fabs in order to reduce costs. Bulk gases are divided into inert, reducing, and oxidizing gases (see Table 5.6 on page 104).

TABLE 5.6 Bulk Gases

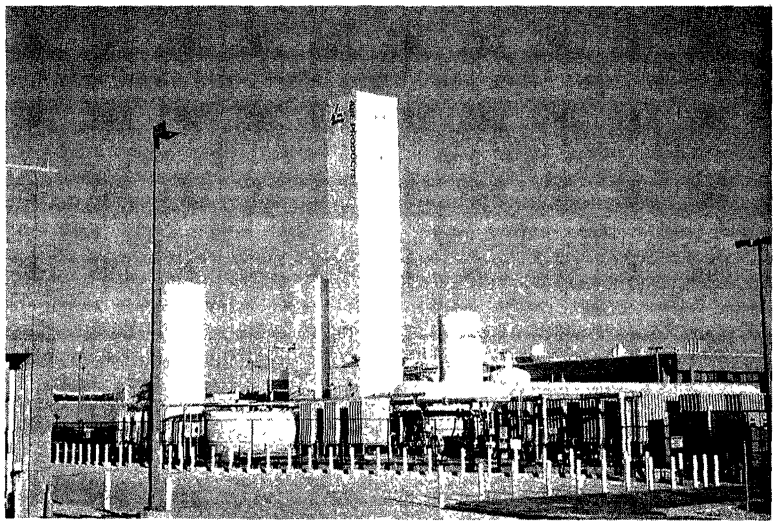
Type of Gas	Gas	Symbol	Example of Use
Inert	Nitrogen	N_2	Purge gas lines and process chambers of moisture and residual gas. N_2 is sometimes used as a process gas during some deposition processes.
	Argon	Ar	Used in process chambers during wafer processing.
	Helium	He	Process-chamber gas also used to leak check vacuum chambers.
Reducing	Hydrogen	H_2	Carrier gas for epitaxial layer process. Also used in combination with O_2 to form steam during some furnace oxidation processes. Used in many wafer fabrication processes.
Oxidizing	Oxygen	O_2	Process-chamber gas.

In recent years, safety concerns have driven the development of on-site generation for other gases, in particular arsine and phosphine.[9] This arrangement permits these highly toxic gases to be produced close to the process tool.

Specialty Gases ■ Specialty gases are those gases supplied in relatively low volumes. These gases are frequently more hazardous than the bulk gases. The specialty gases are the source of many of the materials needed to manufacture a microchip. The greatest challenge for dealing with specialty gases is that many of them are a chemical hazard. They are corrosive (e.g., HCl and Cl_2), pyrophoric (e.g., silane), toxic (e.g., arsine and phosphine), and extremely reactive (e.g., WF_6). The specialty gases are typically used in process chambers at the workstation tools.

Specialty gases traditionally have been transported to the wafer fab in 100-lb metal containers called cylinders for distribution of the gas to the workstations. The cylinders are placed in a gas cylinder cabinet containing a control panel with regulators to control pressure, a flow control, shut-off valves, and a purge panel to control purge sequences needed between cylinder changes. The cabinet also may contain filters, purifiers, or equipment to monitor gas purity, as well as safety equipment such as fire sensors and leak-detection sensors.

A local gas distribution system in the process area is used to deliver the specialty gas from the cylinder to the work chamber of the process tool (see Figure 5.15). This piping can cover hundreds of feet with many permanent welds, bends and fittings and is connected to the process tool at the tool "drop."



Bulk Gas Distribution System
(Photo courtesy of Air Products and Chemicals, Inc.)

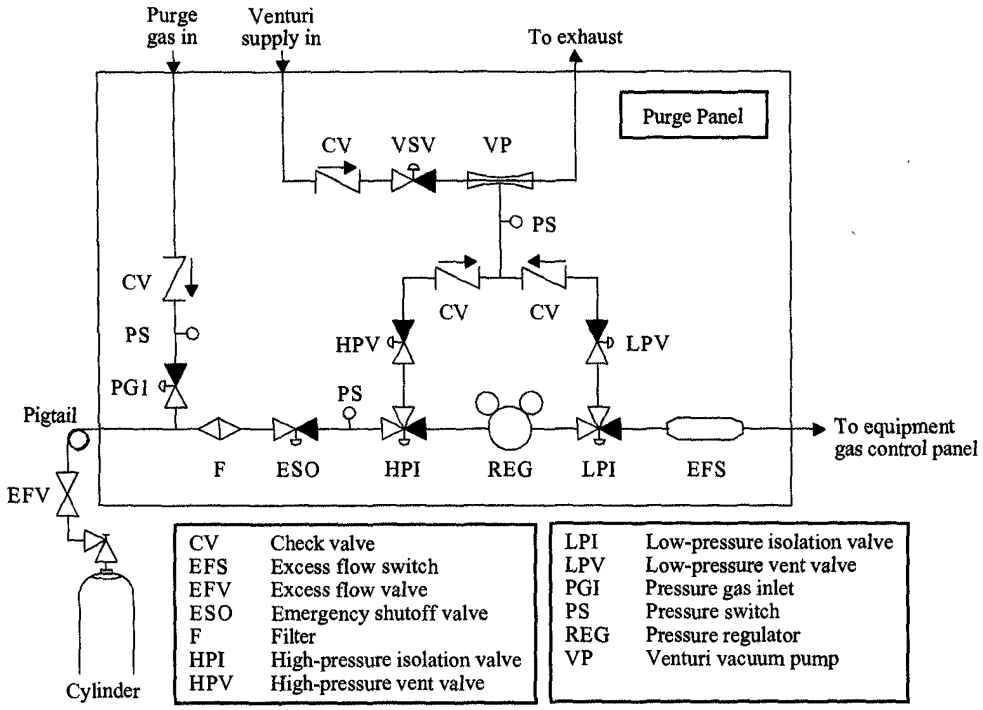


FIGURE 5.15 Typical Specialty Gas System Design
(Schematic used with permission from International SEMATECH)

Gas Purge. A *gas purge* is a method of flushing undesirable residual gases, atmospheric gases, or water vapor from a process chamber and the gas delivery system. Purging can eliminate stagnation points and contamination in gas lines. It is done with an inert gas such as nitrogen and involves replacing the undesirable gas with the purge gas, either through displacement or by pulling the gas out of the system by vacuum flow. Purging of the system on automated equipment is done automatically through software control of gas line valves before and after situations such as a gas cylinder change or opening a process chamber.

Gas Piping. Gas delivery piping is constructed with electropolished 316L stainless steel tubes to transport the gas (316L is a specific type of stainless steel). There are no plastic parts in a gas piping system, except for some membrane gas filters. Double-walled tubing is often used for hazardous gases (see Figure 5.16). The inner walls of the double-walled tubing are electropolished to minimize contamination. *Electropolishing* is a chemical process done to remove about 30 microns of the pipe's inner surface, creating a clean, smooth surface that minimizes the possibility of chemical reactions that produce contaminants. An electropolished surface finish brings a thin layer of chromium to the steel's surface, which emits very few particles.[10]

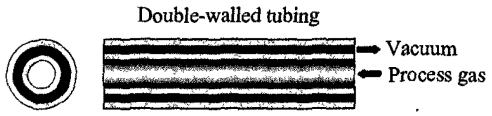


FIGURE 5.16 Double-Walled Tubing
(Used with permission from International SEMATECH.)

Gas Line Connections. For safety reasons, gas lines will have a 360-degree bend, referred to as a pigtail, to make the line more flexible. The gas line terminates in a CGA (Compressed Gas Association) connector for connection to the gas cylinder valve (see Figure 5.17 on page 106). A recently developed DISS (Diameter Index Safety System) cylinder valve has been introduced with specific connector diameters for different gases, which creates a better seal with less abrasion to reduce the potential for contamination. DISS valves are generally used with high-purity products, toxics, and corrosives. Gas lines need a continuous flow through the piping system with no dead spots (known as a deadleg) to trap gas and cause undesirable chemical degradation due to stagnation.

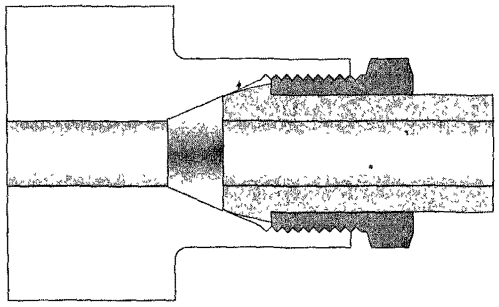


FIGURE 5.17 CGA Gas Line Connector
(Used with permission from International SEMATECH)

Gas Stick. The gas line from the local gas distribution system connects into another gas panel at the process tool. The tool panel is made up of a series of gas "sticks," each controlling one type of incoming gas. Each stick will typically have an on/off valve, flow controller, pressure regulator, and filters (see Figure 5.18). The number of incoming gas lines depends on the process, with a typical number being from six to 30 or more lines for multi-chamber process tools.

Cylinder Change-Out. Cylinders for specialty gases require technicians to change them when they are empty. This is referred to as a *cylinder change-out* and is sometimes done as often as several times per week. Because specialty gases are often hazardous, gas cylinder change-outs require care to avoid safety and product problems.[11] One safety concern during a cylinder change is incorrectly purging gas lines, which causes residual gas to escape, leading to a flame or vapor cloud. Another safety concern is an improperly supported cylinder that may fall over, causing a leak. With regard to the product performance, an incorrect cylinder change can lead to contamination in the gas line and affect product yield. Furthermore, a cylinder change may require the process chamber to be requalified to ensure the gas is acceptable, which interrupts manufacturing production.

Due to safety hazards, contamination concerns, and high maintenance costs, there has been an effort to convert some specialty gases to bulk distribution. For example, a tube trailer of silane is estimated to eliminate the change-out of 300 to 400 cylinders per year, while the tube trailer is changed only once per year.[12] This is a substantial reduction in risk for a problem during change-outs. Eventually more bulk specialty gas distribution systems similar to those used for wet chemicals and bulk gases may replace the cylinders commonly used to deliver specialty gases.[13] This

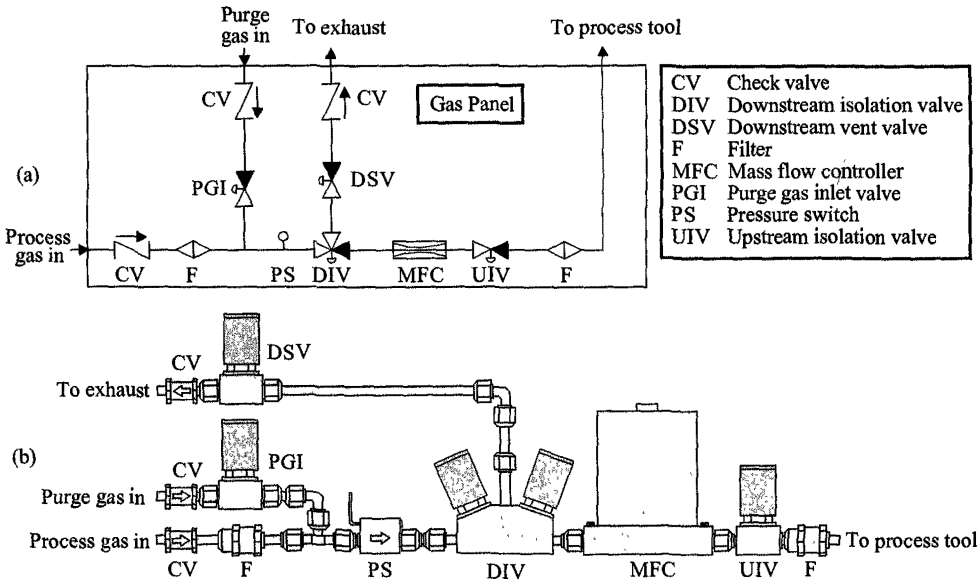
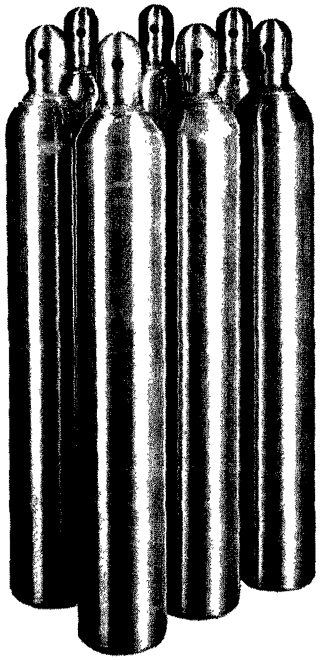


FIGURE 5.18 Gas Stick at Process Tool.
(Schematic (a) used with permission from International SEMATECH)
Component diagram (b) redrawn from Swagelok components, *Swagelok Co. Catalog* provided by Arthur Valve & Fitting Co., Austin, TX.

trend initially started with silane because it is pyrophoric (it burns upon exposure to air) and now encompasses other specialty gases used in high volume, such as nitric oxide (NO). The goal is to reduce the human involvement in the gas delivery system and the potential for error.



Specialty Gas Cylinder
(Photo courtesy of Praxair)

Classifying Specialty Gases ■ Specialty gases may be classified as hydrides (or similar gases), fluorinated compounds, or acid gases. Hydride gases contain hydrogen, while fluorinated compounds contain fluorine. Some common specialty gases are listed in Table 5.7 on page 108, with an example of their use.

TABLE 5.7 Some Common Specialty Gases for Semiconductor Manufacturing

Class of Gas	Gas	Symbol	Example of Use
Hydrides	Silane	SiH_4	Source of silicon in deposition processes (Chapter 11)
	Arsine	AsH_3	Source of arsenic for n-type doping of silicon wafers (Chapter 17)
	Phosphine	PH_3	Source of phosphorous for n-type doping of silicon wafers (Chapter 17)
	Diborane	B_2H_6	Source of boron for p-type doping of silicon wafers (Chapter 17)
	Tetraethyl orthosilicate (TEOS)	$Si(OC_2H_5)$	Source of silicon dioxide used in deposition processes (Chapter11)
	Silicon tetrachloride (also tetrachlorosilane)	$SiCl_4$	Source of silicon used in deposition processes (Chapter 11)
	Dichlorosilane (DCS)	SiH_2Cl_2	Source of silicon used in deposition processes (Chapter 11)
Fluorinated Compounds	Nitrogen trifluoride	NF_3	Source of fluoride icons used in plasma etch processes (Chapter 16)
	Tungsten hexafluoride	WF_6	Source of tungsten used in metal deposition processes (Chapter 12)
	Tetrafluoromethane	C_2F_4	Source of fluoride ions used in plasma etch processes (Chapter 16)
	Carbon tetrafluoride	CF_4	Source of fluoride ions used in plasma etch processes (Chapter 16)
	Silicon tetrafluoride	SiF_4	Source of silicon and fluorine ions for use in deposition, implant, and etch processes (Chapters 11, 16, 17)
	Chlorine trifluoride	ClF_3	Process chamber cleaning gas
Acid Gases	Boron trifluoride	BF_3	Source of boron for p-type doping of silicon wafers (Chapter 17)
	Chlorine	Cl_2	Source of chlorine used in metal etching (Chapter 16)
	Boron trichloride	BCl_3	Source of boron for p-type doping of silicon and a source of chlorine for metal etching (Chapters 16, 17)
	Hydrogen chloride	HCl	Process chamber cleaning gas and contamination getterer
Other Gases	Ammonia	NH_3	Process gas used with SiH_2Cl_2 for deposition of SiN_3 (Chapter 11)
	Nitrous oxide	N_2O	Source of oxygen for reaction with silicon to from silicon dioxide (Chapter 11)
	Carbon monoxide	CO	Used in etch processes (Chapter 16)

SUMMARY

There are four states of matter: solid, liquid, gas, and plasma. Important chemical properties for wafer fabrication are: temperature, pressure and vacuum, condensation, vapor pressure, sublimation and deposition, density, surface tension, thermal expansion, and stress. Process chemicals are widely used to manufacture microchips. Liquids are grouped as acid, base, or solvent. An acid or a base varies in strength and is measured on the pH scale. Chemicals are delivered to the workstation through bulk chemical distribution or point-of-use delivery. Gases used in the wafer fab are classified as either bulk or specialty. Bulk gases are the relatively simple gases used to manufacture chips, such as nitrogen and oxygen. Specialty gases are often a chemical hazard and are transported and stored in the fab in metal cylinders. There are special procedures and delivery systems to ensure that the specialty gases are cleanly and safely used at the workstations. Specialty gases are generally classified as either hydrides, fluorinated compounds, or acid gases.

KEY TERMS

process chemicals
solid
liquid
gas
inert gas
plasma
properties
physical properties
chemical properties
chemical reaction
temperature
heat (thermal energy)
Celsius
Kelvin
pressure
vacuum
condensation
vaporization
absorption
adsorption
vapor pressure
volatile
sublimation
deposition
density
specific gravity
surface tension
coefficient of thermal expansion (CTE)
stress
chemical solution
solute
aqueous solutions
parts per million (ppm)
parts per billion (ppb)
parts per trillion (ppt)
acids
base
pH scale
solvent
bulk chemical distribution (BCD)
ultrahigh purity (UHP)
bulk gases
specialty gases
gas purge
electropolishing
cylinder change-out

REVIEW QUESTIONS

1. What are the four states of matter? Describe each state.
2. What is a material property?
3. Describe the two types of properties of materials.
4. What happens in a chemical reaction? Give an example of a chemical reaction.
5. What is temperature? How is temperature related to heat?
6. State the three temperature scales. Which scale is most common for scientific work?
7. What is pressure? The pressure of a gas is dependent on what conditions?
8. Define vacuum. What is the most common vacuum unit and how is it defined?
9. Define condensation and vaporization. What is the difference between absorption and adsorption?
10. What is vapor pressure? Describe a volatile material.
11. Define sublimation and deposition.
12. Define and describe density. What is the specific gravity of a material?
13. What is surface tension?
14. Define the coefficient of thermal expansion (CTE) of a material.
15. Describe stress and state its formula.
16. How are stress and coefficient of thermal expansion related?
17. What is a chemical solution and what are its components
18. What is an aqueous solution?
19. What do the abbreviations *ppm, ppb,* and *ppt* represent?
20. What is an acid? List three common acids used in a wafer fab.
21. What is a base? List three common bases used in a wafer fab.
22. Explain the pH scale and how it measures acids and bases. Identify the regions in the pH scale attributed to weak solutions of a strong acid, a strong base, and water.
23. What is a solvent? List three common solvents used in a wafer fab.
24. Describe deionized (DI) water for the wafer fab.
25. Describe the bulk chemical distribution system.
26. What is the point-of-use chemical delivery?
27. What are the two types of gas categories?
28. Describe the purity required for gases in semiconductor manufacturing.
29. What five gases are classified as bulk? Give three benefits of bulk gas distribution.
30. What is a specialty gas?
31. What is the greatest challenge for dealing with specialty gases?
32. How are specialty gases usually transported and stored in the wafer fab?
33. What is a gas purge? What type of gas is this done with?
34. How is gas piping constructed? What is electropolishing, and why is this process beneficial?
35. What is CGA? What is DISS and why is it used?
36. Describe a gas stick.
37. Describe a specialty gas cylinder change-out. What are the safety concerns associated with this procedure?
38. State the three classes of specialty gases. Give an example of a common gas for each class.

CHEMICAL SUPPLIERS' WEB SITES

AERONIX Inc.	http://www.aeronex.com
Air Products and Chemicals	http://www.airproducts.com/
Ashland Specialty Co.	http://www.ashchem.com/
ATMI Inc.	http://www.atmi.com/
BOC Edwards	http://www.boc.com/edwards/
Dow Chemical/Filmtec	http://www.dow.com/liquidseps/
Dow Corning	http://www.dowcorning.com/
DuPont	http://www.dupont.com/semiconductor
Eastman Chemical Co.	http://www.eastman.com/
EKC Technology Inc.	http://www.ekctech.com/
J. T. Baker	http://www.jtbaker.com/
Leybold Inficon Inc.	http://www.leyboldinficon.com/
Linde	http://www.linde.de/english/Home.htm
Matheson Gas Products	http://www.mathesongas.com/
Millipore Corp.	http://www.millipore.com/
MKS Instruments Inc.	http://www.mksinst.com/
Parker Hannifin Corp.	http://www.parker.com/
Praxair	http://www.praxair.com/

PTI Advanced Filtration Inc.	http://www.pti-afi.com/
Scott Specialty Gases	http://www.scottgas.com/
Solkatronic Chemicals	http://www.solkatronic.com/
Swagelok Company	http://www.swagelok.com/
Union Carbide	http://www.unioncarbide.com/
Voltaix Inc.	http://www.voltaix.com/

REFERENCES

1. M. Silberberg, *Chemistry: The Molecular Nature of Matter and Change,* (St. Louis: Mosby, 1996), p. 23.
2. W. Runyan and T. Shaffner, *Semiconductor Measurements and Instrumentation,* 2nd ed., (New York, NY, McGraw-Hill, 1997), p. 206.
3. T. Hattori, "Trends in Wafer Cleaning Technology," *Solid State Technology* (May 1995): p. 7.
4. A. Braun, "PPT – Time for a Reality Check?" *Semiconductor International* (June 1998).
5. J. Sargent, V. Starov, and R. Werner, "Transition in the Post-Etch Wafer-Cleaning Market and Technologies," *Solid State Technology* (May 1997): p. 180.
6. R. DeJule, "Bulk Chemical Distribution Addresses Tightening Specs," *Semiconductor International* (August 1996): p. 75.
7. P. Singer, "Effective Gas Handling: A Balance of Cost and Purity," *Semiconductor International* (September 1994): p. 64.
8. P. Singer, "Trends in Gas Management and Use," *Semiconductor International* (April 1998): p. 112.
9. Ibid.
10. A. Braun, "Cleanroom Technologies Continue to Keep Contamination at Bay," *Semiconductor International* (March 1998): p. 59.
11. L. Laurin, "Bulk Silane—a Potential Hazard or a Potential Hazard Reducer," *Solid State Technology,* (January 1998): p. 104.
12. P. Singer, "Trends," p. 112.
13. N. Chowdhury and L. Mostowy, "Developing a Bulk Distribution System for High-Purity Hydrogen Chloride," *Micro Contamination Identification, Analysis and Control* (September 1995): p. 33.

CHAPTER 6

CONTAMINATION CONTROL IN WAFER FABS

Contamination-free wafer fabrication is absolutely essential in order for devices on chips to function properly. As device critical dimensions decrease, contamination requirements become more stringent. We will learn in this chapter about the different types of contamination important for wafer fabrication, their sources, and how to control contamination to effectively manufacture high-performance ICs with minimal contamination-generated defects.

To control unacceptable contamination during manufacturing, the semiconductor industry has developed cleanrooms. A cleanroom is essentially a purified space with ultraclean air that isolates the chip manufacturing from the dirty conditions of the outside world, including chemicals, humans, and ordinary work conditions.

It is important to understand the conditions of a cleanroom because this is where the intricate microchips are manufactured. Numerous details of the work procedures in a wafer fab are defined by how we maintain the integrity of the cleanroom. One of the primary elements for successful cleanrooms is human discipline.

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. State and describe the five different types of cleanroom contamination, and discuss the problems associated with each type of contamination.
2. List seven sources of contamination in a cleanroom, and describe how each one affects wafer cleanliness.
3. Interpret and use the class number for cleanroom air quality.
4. State and discuss seven appropriate actions for workers entering a cleanroom that follow acceptable protocol.
5. Describe the different aspects of an ultraclean cleanroom facility, including air filtering, electrostatic discharge, ultrapure DI water, and process gases.
6. Explain how modern workstation design and a minienvironment contribute to contamination reduction.
7. State the chemistry of the two standard wet-cleaning methods, explain the type of contamination removed by each, and discuss wet-cleaning modifications and alternatives.
8. Describe the different types of wet-cleaning equipment, and state how each cleaning process contributes to wafer cleanliness.

INTRODUCTION

A wafer has multiple microchips on its surface, and each chip has literally millions of devices and interconnection circuitry that are highly sensitive to contamination. As the feature size on a chip shrinks to accommodate higher performance and denser circuitry, the need to control surface contamination becomes more critical (see Figure 6.1 on page 114). To achieve contamination control, all wafer fabrication is done in a cleanroom where contaminants are strictly controlled.

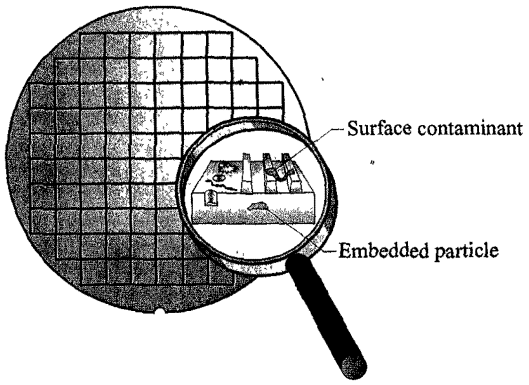
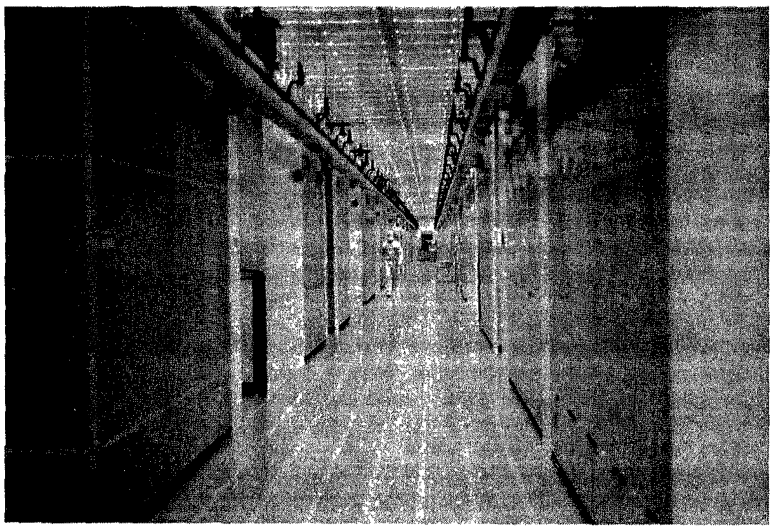


FIGURE 6.1 Wafer Contaminants

Cleanroom Background

When semiconductor manufacturing began nearly half a century ago, the need to control contamination was evident. Early cleanrooms were based around local clean zones, with clean benches used by operators wearing smocks and gloves. The introduction of the high-efficiency particulate air filter (HEPA filter) in the 1960s was a first step toward significant particulate reduction in the wafer fab. The HEPA filters delivered clean air at the workbench to efficiently move particles away from the product.

Modern semiconductor manufacturing is performed in a sophisticated facility known as a *cleanroom*. This is a wafer fabrication facility that is isolated from the outside environment and free from contaminants such as particles, metals, organic molecules, and electrostatic discharge (ESD). When we say free, that means that these contaminants are not detectable at the detection level of the most advanced test instrument. A cleanroom represents the comprehensive set of procedures and practices that are followed to ensure a wafer fabrication facility is contamination-free for semiconductor manufacturing.



Wafer Fab Cleanroom (Photo courtesy of Advanced Micro Devices)

TYPES OF CONTAMINATION

Contamination in semiconductor manufacturing is any undesirable substance introduced to the semiconductor wafer that impacts the production yield or electrical performance of a microchip. Since our interest is wafer manufacturing, we will focus on the various types of surface contamination introduced during the manufacturing process.

Contamination often leads to a defective chip. *Killer defects* are those causes of failure where the chip on the wafer fails during electrical test. It is estimated that 80% of all chip failures at electrical test are due to killer defects from contamination.[1] Failure at electrical test results in a yield loss, causing the defective die on the wafer to be scrapped (thrown away) at a significant cost to the chip manufacturer.

Cleanroom contamination is grouped into five categories:

- Particles
- Metallic impurities
- Organic contamination
- Native oxides
- Electrostatic discharge (ESD)

Particles

Particles are small objects that can adhere to the surface of a wafer. Airborne particles suspended in the air are referred to as *aerosols*. The relative size distribution of various particles from pebbles down to atoms is shown in Figure 6.2.[2]

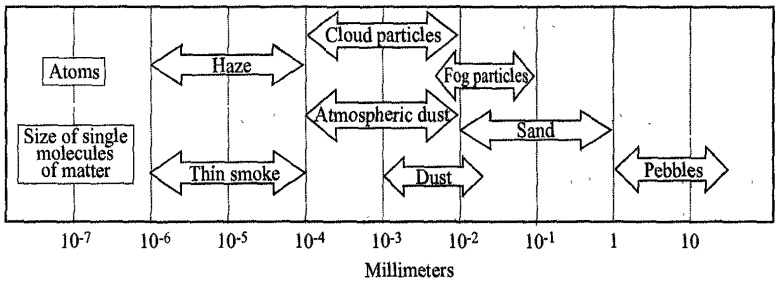
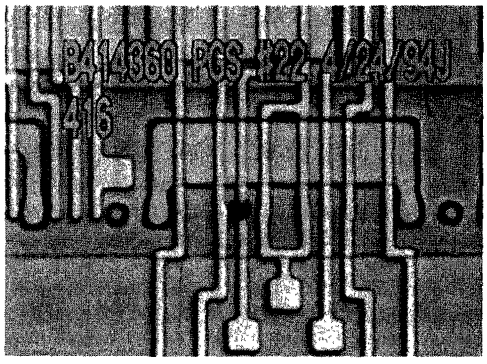


FIGURE 6.2 Relative Size of Particles

Problems From Particles ■ For semiconductor manufacturing, our objective is to control and reduce wafer exposure to particles. Particles can cause defects during the wafer fabrication process leading to open or bridged circuitry. They can create an electrical short between adjacent conductors. Particles also can be sources of other types of contamination, as discussed in the following section.



Defect from Particles (Photo courtesy of Advanced Micro Devices)

The rule of thumb for an acceptable particle size in semiconductor manufacturing is that it must be less than one-half the minimum device feature size.[3] Particles larger than this size will cause killer defects. For example, a 0.18-μm feature size cannot be exposed to 0.09 μm and larger particles. To

appreciate these dimensions, consider that a human hair is about 90 μm in diameter. A dimension of 0.18 μm would be about 500 times smaller than the human hair (see Figure 6.3).

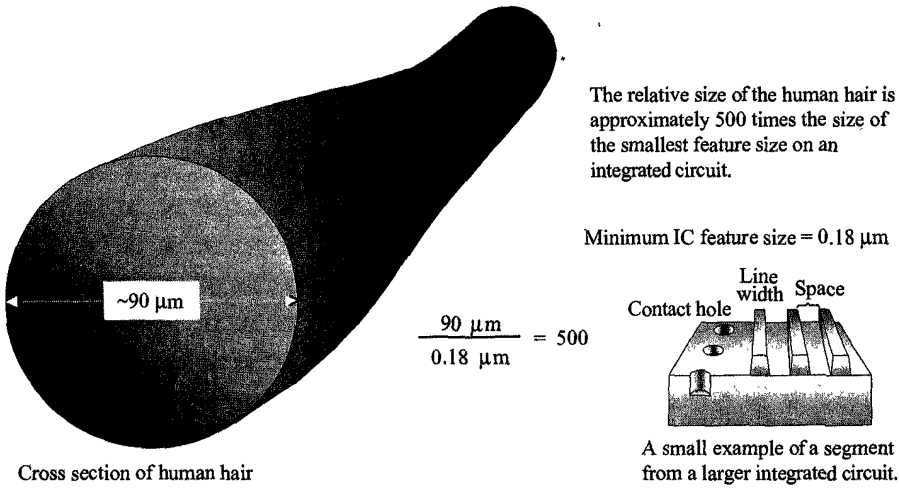


FIGURE 6.3 Relative Size of Human Hair to 0.18 μm Particle Size

The *particle density* on a wafer surface represents the number of particles in a given area. A higher particle density produces a higher chance of a killer defect. The number of particles above a certain critical size that are added to a wafer surface at an operation is termed the *particles per wafer per pass* (*PWP*). In the early days of semiconductor manufacturing, skilled operators visually inspected wafers for particles using simple tools such as a microscope. However, this approach is unacceptable for the VLSI and ULSI eras. Since the mid-1980s, particle detection has been widely done by scanning the wafer surface with a laser beam and detecting the position and intensity of the scattered light caused by particles (see Figure 6.4). The smallest detectable diameter of current particle-detection equipment used in production is roughly 0.1 μm.[4]

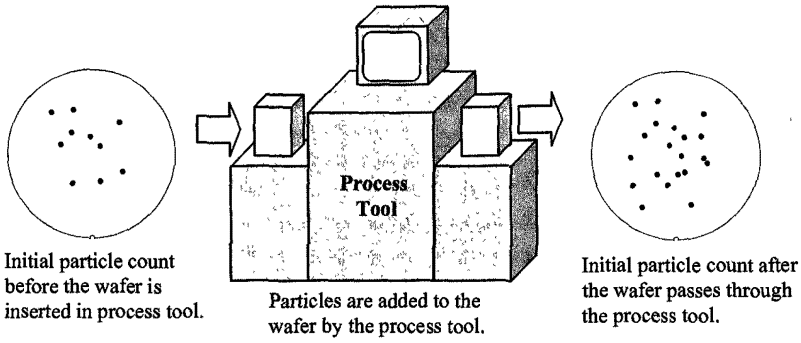


FIGURE 6.4 Particles Per Wafer Per Pass on a Wafer

Semiconductor manufacturing in the ULSI era also requires stringent control of a wide range of airborne molecular particles, including water vapor, acid vapors, hydrocarbons, and gases that add impurities to the wafer surface. The measurement of airborne molecular contamination at the ppb level is now becoming standard for advanced fabrication facilities. Molecular compounds include metals, nonmetals, and organic and inorganic contaminants, all of which are damaging for device performance.

Metallic Impurities

Contamination in a wafer fab also can come from metal compounds. The typical *metal impurities* damaging to semiconductor processing are the alkali metals that are found in many common chemicals and processes. These metals are strictly controlled in all materials used in a wafer fab (see

Table 6.1).[5] The alkali metals are from Group IA of the periodic table and are extremely reactive elements because they give up one valence shell electron, thus forming a cation that reacts with a nonmetal anion to form an ionic compound.

TABLE 6.1 Typical Metal Impurity Elements

Heavy Metals	Alkali Metals
Iron (Fe)	Sodium (Na)
Copper (Cu)	Potassium (K)
Aluminum (Al)	Lithium (Li)
Chromium (Cr)	
Tungsten (W)	
Titanium (Ti)	

Metals come from chemical solutions or different process steps during semiconductor manufacturing. The process of ion implantation (discussed in a later chapter) exhibits the highest metal contamination, on the order of 10^{12} to 10^{13} atom/cm^2.[6] Another source of metallic contamination is through reaction of the chemicals with the transport piping and containers. For instance, carbon monoxide gas, used as an additive gas to improve various wafer processes, can react with the nickel in stainless steel, gaskets, and other components of the gas delivery system.[7] From this reaction gas-phase carbon monoxide forms nickel tetracarbonyl particles that are then distributed on the wafer surface. These particles can be redistributed into the bulk wafer and lead to increased device defects.

There are two basic ways for metallics to deposit on a wafer surface.[8] In the first way, metal binds to the silicon surface by the charge exchange between a metallic ion and a hydrogen atom located on the wafer surface. These types of metallic impurities are difficult to remove. The second way metal is deposited on the wafer surface is when the surface oxidizes and metallic impurities are located in the oxidation layer. For this reason cleanliness is critical in oxidation processes (see Chapter 10). Metallic impurities in an oxide layer are removed only by removing the oxide from the wafer surface.

Metallic ions are highly mobile in semiconductor materials and are referred to as *mobile ionic contaminants* (*MICs*). When introduced into a wafer, MICs move throughout the silicon and seriously damage the device's electrical performance and long-term reliability. Sodium is typically one of the most prevalent MICs in untreated chemicals, with people as its greatest carrier. The human body contains a high percentage of sodium in the form of fluids (e.g., saliva, tears, perspiration, and so on). Sodium contamination is rigorously controlled in wafer fabs.

It is hard to imagine how mobile an MIC material is in silicon. Consider an MIC such as sodium. A single crystal of table salt (NaCl) contains enough sodium to deposit one quadrillion (10^{12}) atoms of sodium per square centimeter on 5,000 wafers (150-mm diameter) and therefore destroy all chips on the wafers.

Problems From Metallic Impurities ■ Metallic impurities lead to reduced device yield in semiconductor manufacturing, including structural defects in the oxide-polysilicon gate structure.[9] Additional problems include increased leakage currents at a pn junction and reduced minority carrier lifetime. MIC contamination can migrate to the oxide-silicon interface in the gate structure and alter the threshold voltage required to turn on a transistor (see Figure 6.5 on page 118). Because they are so mobile, metallic ions can move around a device long after electrical test and shipment and cause the device to fail during usage. A major goal of semiconductor manufacturing is to minimize exposure to metallic impurities and MICs.

Organic Contamination

Organic contaminants are those that contain carbon, nearly always bonded to itself and to hydrogen, and sometimes to other elements as well. Some sources of organic contamination are bacteria, lubricants, vapors, detergents, solvents and moisture. Equipment used in wafer fabs today is designed with components that require no lubricant, for example, oil-free pumps and bearings.

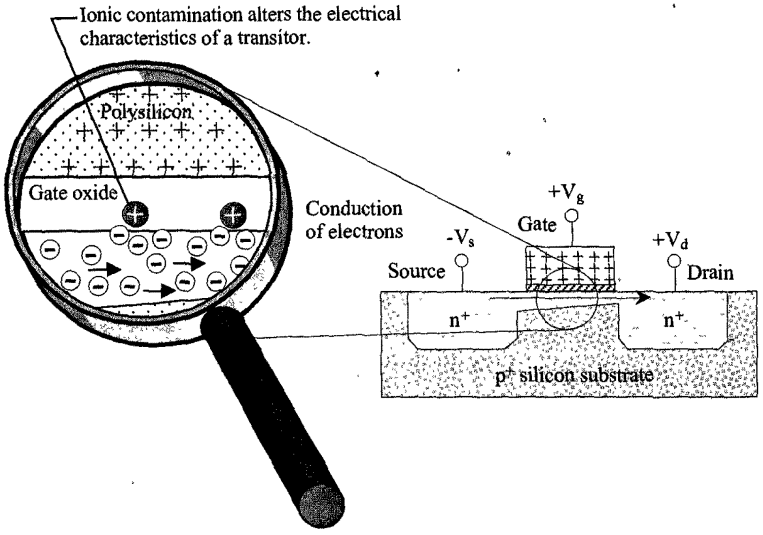


FIGURE 6.5 Mobile Ionic Contaminant Altering Threshold Voltage

Problems From Organic Contaminants ■ Trace organic contaminants degrade the integrity of the gate oxide material under certain processing conditions.[10] Another problem with organic materials on the surface of semiconductor wafers during processing is incomplete cleaning of the surface. This condition permits contaminants such as metal impurities to remain intact on the wafer surface after cleaning.

Native Oxides

The surface of a silicon wafer will oxidize if exposed either to air at room temperature or to DI water that contains dissolved oxygen. This thin oxide layer is termed *native oxide.* The initial native-oxide growth on a silicon wafer occurs in the presence of moisture. When the wafer surface is exposed to air, several tens of molecular layers of moisture adsorb on the wafer within a second. Oxygen from the room is dissolved into the adsorbed moisture layer on the wafer surface and penetrates into the silicon surface. This process causes the silicon surface to oxidize even at room temperature. The thickness of the native oxide increases as exposure time lengthens.

Problems From Native Oxides ■ A silicon surface that is native oxide-free is important for semiconductor performance and reliability. Native oxide interferes with other process steps, such as the growth of single-crystal film on the wafer or the growth of the ultrathin gate oxide.[11] Native oxide also includes some metallic impurities, which can move into the silicon wafer and cause electrical defects.[12]

Another problem created by native-oxide growth occurs in contact regions for metal conductors. Contacts make electrical connections between interconnect wiring and the source and drain region of the semiconductor device. If there is a native-oxide layer, then this will increase the resistance of the contact and reduce and perhaps even prevent current flow (see Figure 6.6).[13]

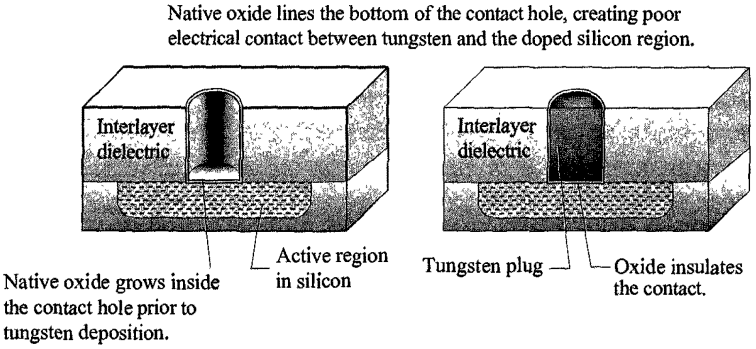


FIGURE 6.6 Native Oxide

Native oxide requires removal through a cleaning step with a mixture of HF acid (see the section on wet cleaning later in this chapter). Another approach to inhibiting native oxide is to integrate multiple process steps into a multichamber tool with an evacuated, high-vacuum chamber so that the wafers are not exposed to ambient atmosphere and moisture.

Electrostatic Discharge

Electrostatic discharge (*ESD*) is a form of contamination because it is an uncontrolled transfer of static charge from one object to another that potentially damages the microchip. ESD is generated when two materials with differing static-charge potential are touching or rubbing each other (this action is referred to as *triboelectricity*). Atoms with excess negative charges are attracted to adjacent atoms that have an excess of positive charges. This attraction creates a discharge of electricity that can be up to tens of thousands of volts.

Semiconductor manufacturing is especially prone to developing static electrical charges because the wafer fab is maintained at a low humidity, typically 40% ± 10% relative humidity (RH). This condition is conducive to the generation of high levels of static charge.[14] Although increasing the RH would decrease static-charge generation, it would also increase contamination through corrosion, thus making it an impractical approach.

Problems From Electrostatic Discharge ■ Although the amount of static charge transferred during ESD is usually small (on the order of nanocoulombs), the discharge deposits its energy into a very small area of the wafer. An electrostatic discharge that occurs in a few nanoseconds can generate peak currents over 1 amp, literally vaporizing metal conductor lines or punching through oxide layers.[15] This discharge also can be the cause of gate oxide breakdown. Another significant problem with ESD is that once a wafer surface has a charge buildup, its resulting electric field can attract charged particles or polarize and attract neutral particles to its surface (see Figure 6.7). An example of this is how a TV screen attracts dust particles. Furthermore, the smaller the particle, the more effect the electrostatics have on particle attraction.[16] As device critical dimensions are reduced, smaller particles attracted by ESD become more significant and can create killer defects. In order to minimize particle contamination, wafer charging must be controlled.

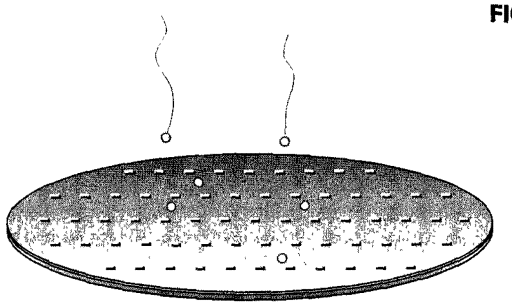


FIGURE 6.7 Particles Attracted to a Charged Wafer

SOURCES AND CONTROL OF CONTAMINATION

The wafer fab cleanroom is stringently controlled for contamination to reduce killer defects that impact microchip performance. Nearly anything that comes in contact with a wafer is a potential source of contamination. Seven sources of contamination in a wafer fabrication facility are:

- Air
- Humans
- Facility
- Water
- Process chemicals
- Process gases
- Production equipment

Air

The most fundamental concept of a cleanroom is the control of particles in the wafer fab air. The air that we normally breathe is not acceptable for semiconductor manufacturing because it contains excessive airborne contaminants. These small airborne aerosol particles float and remain in the air for a long period of time, depositing on wafers as contamination and creating killer defects.

The *class number* designates the air quality inside a cleanroom by defining the particle size and density in the cleanroom air. This figure represents how well particles are controlled to reduce particulate contamination. Class numbers originated from Federal Standard 209, first released in 1963 and revised several times until the most recent version of 209E (see Appendix B). Table 6.2 shows the number of acceptable particles per cubic foot for the different cleanroom class numbers and particle size.

TABLE 6.2 Definition of Airborne Particulate Cleanliness Classes Per Federal Standard 209E

	Particles/ft^3				
Class	**0.1 μm**	**0.2 μm**	**0.3 μm**	**0.5 μm**	**5 μm**
1	3.50×10	7.70	3.00	1.00	
10	3.50×10^2	7.50×10	3.00×10	1.00×10^1	
100		7.50×10^2	3.00×10^2	1.00×10^2	
1,000				1.00×10^3	7.00
10,000				1.00×10^4	7.00×10
100,000				1.00×10^5	7.00×10^2

If the cleanroom class is stated only by the number of particles, such as a cleanroom of class 1, then assume a particle size of 0.5 μm. This means there is a maximum of one particle of size 0.5 μm or larger per cubic foot. For particle sizes different from 0.5 μm, the cleanroom class should be expressed as the class number at a specific particle size. Examples are: class 10 at 0.2 μm (read from Table 6.2 as a maximum of 75 particles per cubic foot at 0.2 μm or larger) and class 10 at 0.1 μm (a maximum of 350 particles per cubic foot at 0.1 μm or larger).

Ultrafine Particles ■ There recently has been usage of a class 0.1, with particle sizes down to 0.02 to 0.03 μm. The latest clean air standard also has a provision for the number of *ultrafine particles* in a cubic meter of air, which is called the "U" descriptor. The U descriptor specifies ultrafine particles as those smaller than 0.1 μm in diameter, down to the smallestdiameter detectable with a discrete particle counter. Without referring to a specific particle size, the U descriptor defines cleanliness as U(x), where (x) is the maximum allowable number of ultrafine particles per cubic meter of air.

Humans

A human is a particle generator. People continually enter and leave cleanrooms and are the greatest sources of contamination in a cleanroom.[17] Particles come from hair and hair products (hair spray, gel), lint from clothes, flakes of dead skin, and so on. On the average, a person sheds over one ounce of particles per day, which can amount to an astonishing 10,000,000 particles per minute of 0.3 μm size and larger (see Table 6.3).

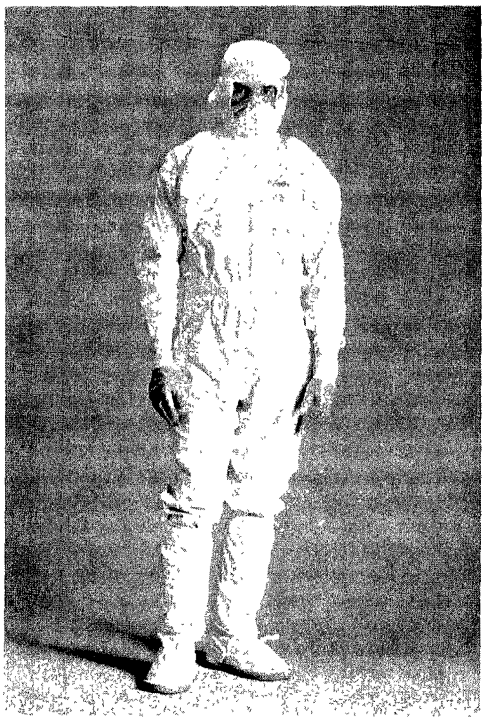
TABLE 6.3 Particles Emitted by Human Activities

Source of Particles	Average Number of Particles per Minute > 0.3 μm
Motionless (sitting or standing)	100,000
Moving hands, arms, trunk, neck, and head	500,000
Walking at 2 miles per hour	5,000,000
Walking at 3.5 miles per hour	7,500,000
Cleanest skin (per square foot)	10,000,000

Simple activities in the wafer fab such as opening and closing doors or excessive movements around process tools create particulate contamination. Normal human activities such as talking, coughing, and sneezing are damaging to semiconductors.

Cleanroom Garments ■ To attain ultraclean conditions in a cleanroom, humans must follow certain procedures, known as *cleanroom protocol,* and be covered with special cleanroom garments (also referred to as a "bunny suit"). The garment is made up of a hood with a facemask, coveralls, boots, and gloves, and will completely cover the body. The goal of the cleanroom garment system is to meet these functional criteria:

- Total containment of body-generated particles and aerosols.
- Zero particle release from the garment system.
- Zero electrical-charge buildup for ESD.
- No release of chemical or biological residues.



Technician in Cleanroom Garment. (Photo courtesy of International SEMATECH)

The modern cleanroom garment is a high-technology membrane fabric or densely woven polyester fabric. Advanced materials will have a 99.999% efficiency rating (stops 99.999% of all particles from passing) for 0.1 micron particles and greater. The extent to which the cleanroom garment system covers the human body can vary. Some fabs require inner clothing layers, such as polyester underwear, in addition to the outer bunny suit. People may be required to take a shower and use lotion to prevent skin flaking prior to final dressing and entering fabs. Some wafer fabs also require fab workers to put on a bubble helmet with a breathing recirculator and exhaust blower that pumps the user's breath through a filter pack strapped to the waist. This helmet prevents particles from human saliva from contaminating the fab work area.

Cleanroom Protocol ■ Each semiconductor company has a strict procedure for cleanroom protocol to minimize contamination in the cleanroom. Some standard cleanroom protocols are listed in Table 6.4.[18]

TABLE 6.4 Proper Cleanroom Protocol

Should Do	Should Not Do	Why?
Only authorized personnel are allowed within the cleanroom.	No people allowed that have not been properly trained in what the cleanroom expects of them. The cleanroom supervisor has the last word.	Authorized personnel are familiar with the many strict and demanding restrictions of cleanroom operations.
Take only what is necessary into the cleanroom.	No cosmetics, tobacco products, handkerchiefs, tissues, food, drink, candy, wooden/mechanical pencils or pens, perfumes, colognes, watches, jewelry, cassette players, phones, beepers, video cameras, audiocassette recorders, gum, combs, hair brushes, cardboard or noncleanroom-approved paper. No blueprints, operations manuals or instruction sheets.	Sources of unwanted contaminants.
Gown in the prescribed manner according to your company training.	No uncovered street clothes allowed within the cleanroom.	The proper cleanroom-approved garmenting protects the product from the contamination associated with humans.
Always make sure that all head and facial hair is covered.	Do not expose any facial or head hair.	Sources of unwanted contaminants.
Follow procedures for entering the cleanroom, such as an air shower (if required).	Do not open any door into the cleanroom until all procedures are complete.	Air showers may assist in removing contaminants; many firms have stopped using this procedure due to problems with airborne contamination.
Keep cleanroom garment closed at all times while in the cleanroom.	Do not expose any street clothing while in the cleanroom. Do not allow any part of your skin to touch the outer parts of the cleanroom garment.	Sources of unwanted contaminants.
Move slowly.	Do not congregate or move quickly.	This disrupts the airflow pattern.

Facility

For semiconductor manufacturing to function as an ultraclean environment, a systems approach is necessary to control all inputs and outputs to the cleanroom area. There are three basic strategies for eliminating particles from cleanrooms:[19]

1. Start out with a cleanroom that is free of particles.
2. Minimize the introduction of particles into the cleanroom through equipment, tools, personnel, and cleanroom supplies.
3. Continuously monitor the cleanroom for particles for timely response to cleaning maintenance.

Cleanroom Layout ■ In the early 1970s, the LSI manufacturing environment had an overall cleanliness class of 10,000 in the manufacturing space and a local class of 100 at the individual workbenches. The industry developed a *ballroom layout* approach, with one large fabrication room with a class 10,000 rating and laminar flow benches that provided the class 100 work environment (see Figure 6.8).

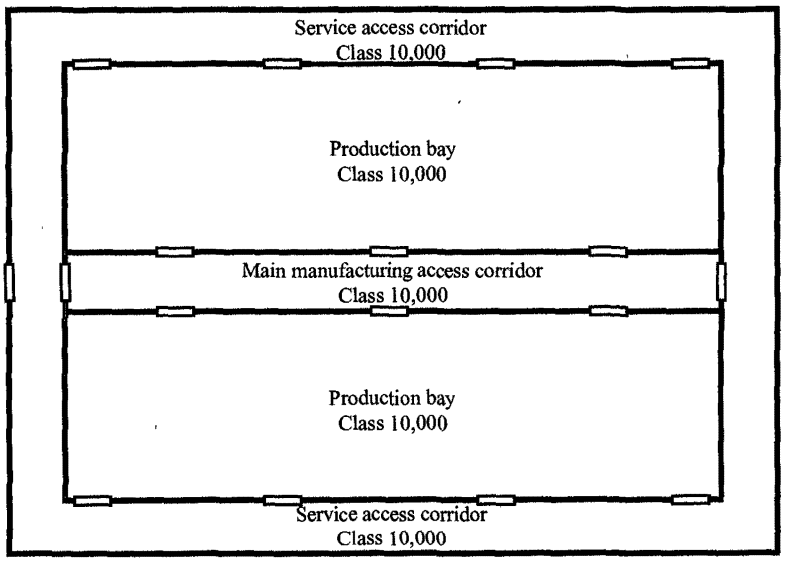


FIGURE 6.8 Early Ballroom Cleanroom Layout

With the submicron technologies of the 1980s came the introduction of the *bay and chase layout* approach to cleanrooms. In this cleanroom layout, a common corridor separates the production area (referred to as production bay or process bay) from the service area (referred to as service chase, equipment chase, or gray area), as shown in Figure 6.9 on page 124. Inside the production bay where wafers are processed, the cleanliness class is typically class 1. Most of the equipment maintenance occurs in the class 1,000 service chase.

Modern wafer fabs are built based on either type of cleanroom design. The ballroom approach is promoted as an advanced cleanroom design that encompasses automated wafer handling and localized contamination control at the process tools.[20] For both types of cleanrooms, there is typically a *sub-fab* area below the cleanroom that contains much of the facilities equipment (e.g., pumps, piping, ductwork, cabling, and so on).

Airflow Principles ■ To achieve ultraclean conditions in a cleanroom, the nature of the airflow is critical. For a cleanroom class of 100 and below, *laminar airflow* is necessary.[21] Laminar airflow means that the airflow is smooth with no turbulence in the airflow pattern (see Figure 6.10 on page 124). The vertical laminar airflow has a slight positive pressure relative to the outside pressure and acts as a curtain to minimize cross-contamination from equipment or personnel to any exposed product.[22]

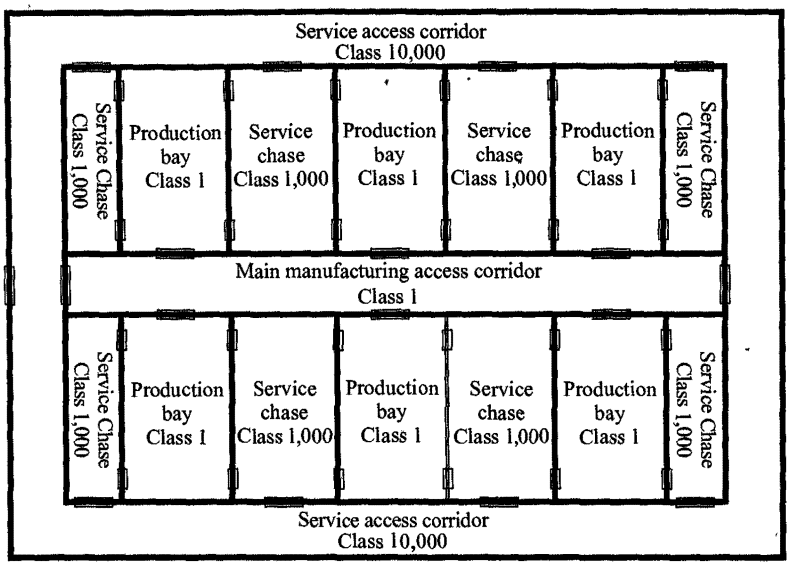


FIGURE 6.9 Bay and Chase Concept for Cleanroom

Air Filtering ■ A simplified diagram of the air handling system in a wafer fab cleanroom is shown in Figure 6.11. Air enters the cleanroom through special particulate filters in the ceiling and passes with laminar flow to the floor and into the recirculation air system to return back to the air filtering system with makeup air. In a modern fab, air may turn over every six seconds to improve recovery of ultraclean conditions during disturbances such as shift changeover. An exhaust system is used to remove undesirable heat and chemicals from the process tools and work areas.

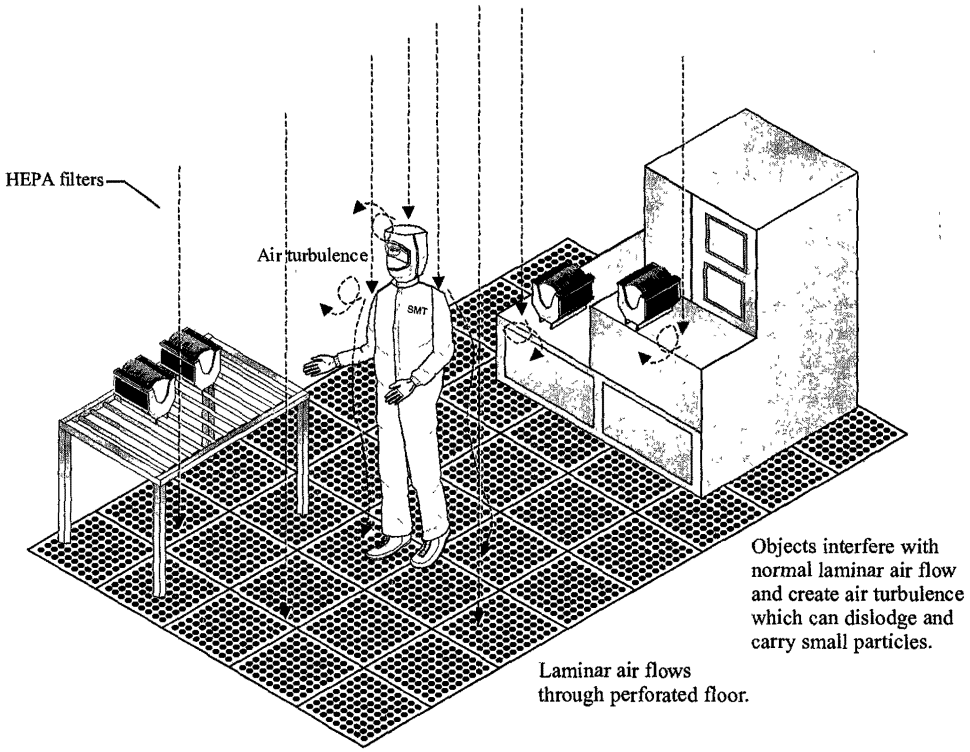


FIGURE 6.10 Laminar Airflow

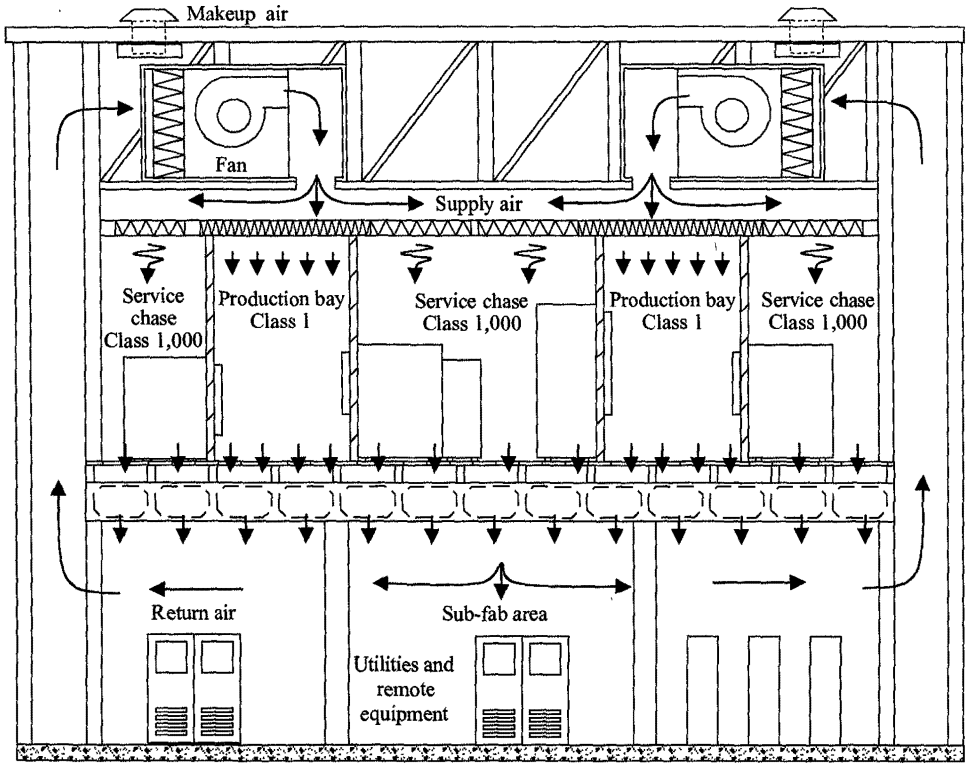


FIGURE 6.11 Wafer Fab Air Handling System

The special particulate filters located in the ceiling are either *high-efficiency particulate air* (*HEPA*) or *ultra-low penetration air* (*ULPA*). HEPA fibrous filters are made with fiberglass fibers and constructed to create laminar airflow (see Figure 6.12). In general, ULPA represents filters that have an efficiency of 99.9995% or better for particulate diameters greater than 0.12 microns.[23]

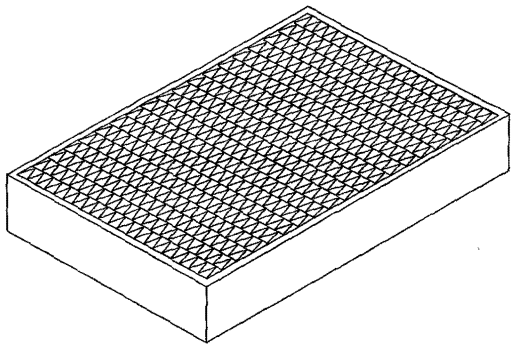


FIGURE 6.12 HEPA Filter

Temperature and Humidity ■ Temperature and humidity settings are specified for wafer fab facilities. An example of temperature control in a class 1 at 0.3 μm cleanroom is 68 ± 0.5°F.[24] Relative humidity (RH) is important because of its contribution toward corrosion (higher RH promotes more corrosion, just as moisture causes steel to rust). A typical RH setting may be 40% +/- 10%.

Electrostatic Discharge ■ Most electrostatic discharge (ESD) problems are controlled through the proper use of equipment and procedures. The principal ESD control methods are:

- Static-dissipative cleanroom materials
- ESD grounding
- Air ionization

Cleanroom materials such as carts, cassettes, equipment, and so on must be *static dissipative*. This term means that the resistivity of the material is lowered through the use of conductive additives, permitting mobile electrostatic charges to flow through the material. At the same time, people and objects in the cleanroom must be continuously connected to ground. In this manner, electrostatic discharge is conducted through the human body and all cleanroom materials that the wafer comes in contact with and flows harmlessly away from the product through ground. This setup avoids potentially discharging through a device on a wafer and causing irreparable damage to the chip.

Air Ionization. The wafer has various insulating materials placed on it during processing, such as an oxide film. These materials are easily charged and will hold this charge for a long period. Since this insulation material is in intimate contact with the product, it requires a method to neutralize the charge buildup. The most common way to neutralize insulating materials on a wafer is with *air ionization.* Special ionizer emitters located on the ceiling in the cleanroom produce a high electric field that ionize the air molecules and make it conductive by gaining or losing an electron. When this conductive air contacts a charged surface, such as an insulator material on a wafer, the surface attracts ions from the opposite polarity and neutralizes the electrostatic surface charge (see Figure 6.13).

Air ionization by emitters can be limited because many ions are eliminated before reaching the wafer surface due to recombination. A recent development is air neutralization using soft-X-ray radiation.[25] Exposing the ambient air surrounding the charged wafer to soft X-rays will generate large ion pairs. This process effectively neutralizes the charge on the wafer surface to 0 volts after roughly two seconds.

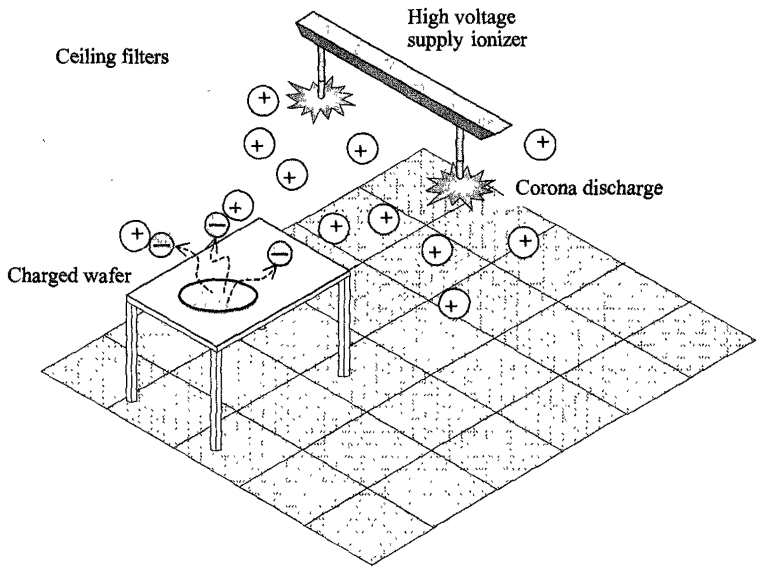


FIGURE 6.13 Neutralizing Static Charge on a Surface with Air Ionization

Water

To manufacture semiconductors, a high quality, *ultrapure deionized* (*DI*) water (sometimes referred to as UPW) is required in large quantities. City water has too many contaminants to be acceptable for the production of wafers. DI water is the most heavily used chemical in semiconductor manufacturing, primarily in the chemical wafer cleaning solutions and as a postclean rinse. It is estimated that current DI water consumption in a wafer fab runs up to 2,000 gallons of ultrapure deionized water for each wafer produced in a modern 200-mm process line.[26]

Unacceptable contaminants in ultrapure DI water are:

- Dissolved ions
- Organic materials
- Particulates
- Bacteria
- Silica
- Dissolved oxygen

Figure 6.14 shows different water particles and their sizes.[27]

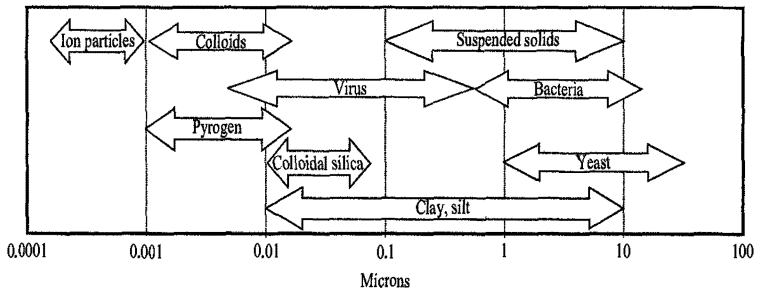


FIGURE 6.14 Size of Particles in Water

Dissolved ions in water are from minerals such as sodium and potassium that easily form ions. An example is salt (NaCl) that breaks up into Na^+ and Cl^- ions. These mobile ionic contaminants (MICs) are undesirable and create performance problems in semiconductor devices.

Organic materials, referred to as *total organic carbon (TOC)*, are the sum of all carbon-containing compounds dissolved in the water. These organic contaminants have a damaging effect on the ability to grow oxide films.

Bacteria are live, reproducing organisms in the water. Fragments shed by bacteria are known as *pyrogens*. Bacteria found in water lead to defects in oxidation, polysilicon, and metal-conducting layers. Some bacteria contain phosphorus that could cause uncontrolled doping.

Silica is found in city water as finely suspended particles. These particles range in size from ten angstroms to ten microns. A high silica content fouls filtration equipment in the water purification equipment and also decreases the reliability of thermally grown oxides.

Another type of contaminant in water is dissolved oxygen. It creates a problem because it leads to native-oxide formation on the wafer surface. Dissolved oxygen in water also creates problems if the water is pressurized as it is in some semiconductor processes. When the water is depressurized, dissolved gases may come out of solution forming bubbles that can cause incomplete water wetting on the wafer surface.[28]

DI-Water Installation ■ A DI-water installation has two major parts for purifying water, referred to as the makeup loop and the polishing loop (see Figure 6.15 on page 128).[29] The *makeup loop* removes particles, total organic carbons (TOCs), bacteria, microorganisms, ionic impurities, and total dissolved minerals from the raw water. It has a prefilter that removes particles above one micron and a purifier section for removal of ionic impurities, bacteria, and dissolved gases. The *polishing loop* is the final part of the water purification system that removes the remaining contaminants.

Deionization ■ *Water deionization* to make DI water is the process of removing the electrically active salt ions using specially manufactured ion-exchange resins. This process changes the water from a conductive medium to a resistive medium with a resistivity of 18 megohm-cm (18,000,000 ohm-cm) at 25°C. DI water used in the wafer fab is referred to as 18-megohm (MΩ) water. Ultrapure DI water is passed through two deionizers, once in the makeup loop and another time in the polishing loop.

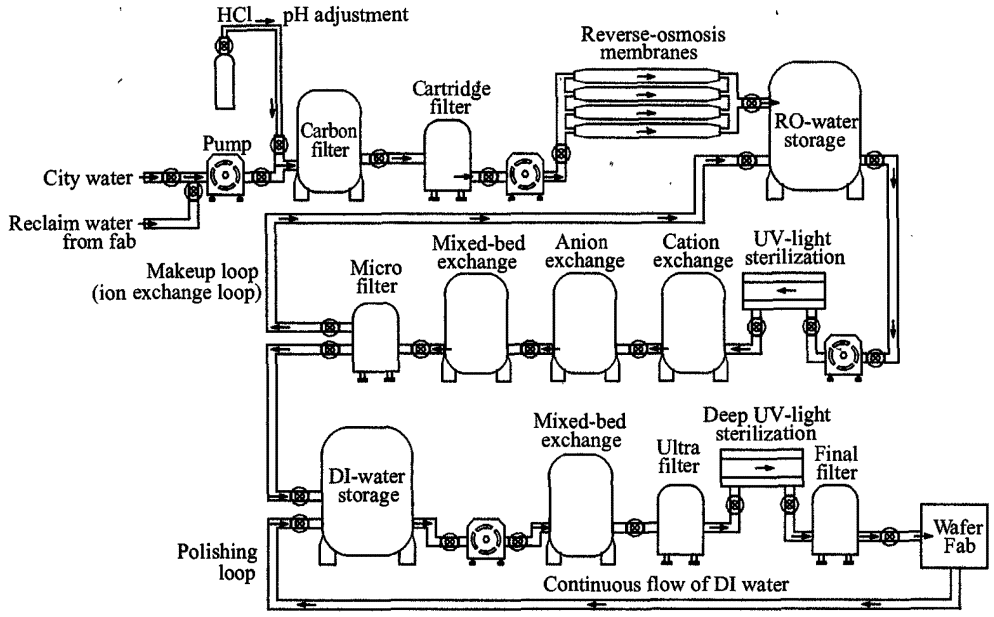


FIGURE 6.15 DI-Water Makeup and Polishing Loops

DI-Water Filtration ■ The DI-water makeup loop employs various filters with the intent of letting water pass through while trapping the particulates and colloids (extremely fine particles) in the filter media. A common filtering technique for ultrapure DI water is *reverse osmosis* (*RO*) to remove smaller particles and metallic ions. The operating principle for RO filters is to flow the water under pressure across a membrane filter to separate ionized salts, colloids, and organic materials down to 150 molecular weight (see Figure 6.16). RO can separate impurities as small as 0.005 microns and is also referred to as *hyperfiltration*.

Ultrafiltration is used at point-of-use (POU), meaning its filtering location is at the equipment, to remove submicron-size particles. This filter uses pressure and flow through a membrane with pore sizes ranging from 10 angstroms to 0.2 microns.

Other parts of an ultrapure DI-water system include a degasifier unit used to remove dissolved gases in DI water, such as oxygen, to reduce contamination such as native-oxide growth on the wafer surface. An improved technique for removing dissolved gases (primarily oxygen) to less than 10 ppb is a *membrane contactor* (see Figure 6.17).[30] Membrane contactor filters consist of

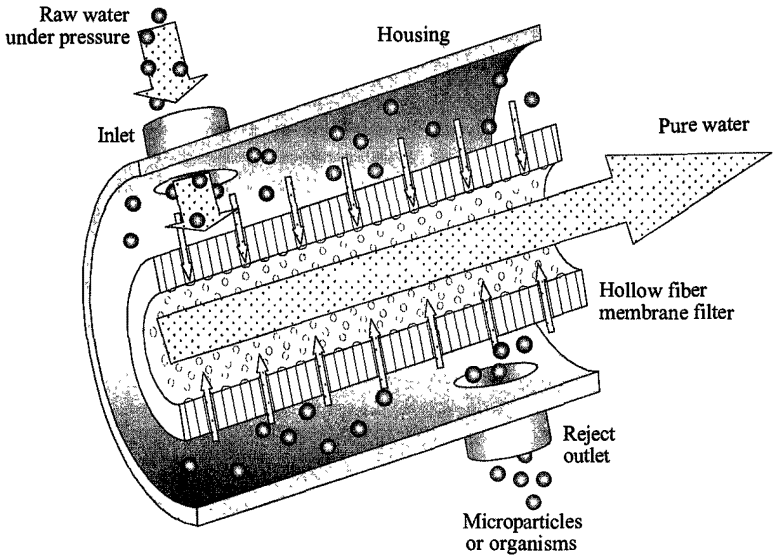


FIGURE 6.16 Principle of Reverse Osmosis Filtration

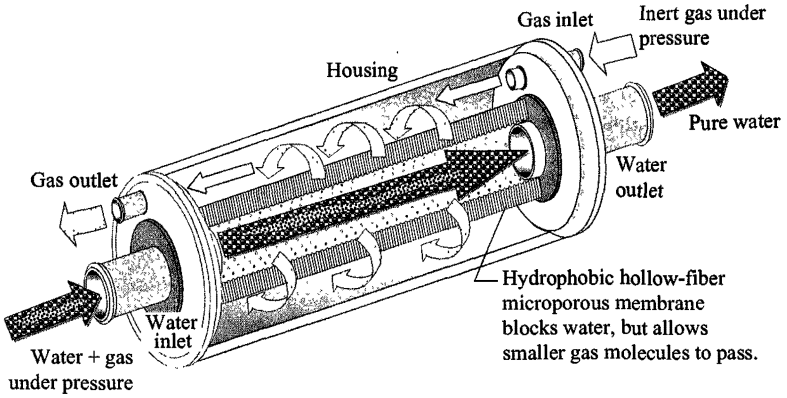


FIGURE 6.17 Membrane Contactor Filter

hydrophobic polypropylene microporous hollow fibers. The fibers have small pores in their walls to allow dissolved gases to pass through. However, the fiber's hydrophobic nature does not permit liquid water to pass through the pores. The membrane removes oxygen and all free gases.

Zeta Potential ■ The *zeta potential* represents a positive or negative electrical charge that can build up in colloids (very fine suspended particles in a liquid). Particles in water generally have a zeta potential with a negative charge, such as colloidal silica, bacteria, and pyrogens. These colloids can be filtered out of the water using a positively charged filter that traps particles that are smaller than the pore diameter of the filter.[31] This form of particle removal uses electrostatics to attract particles with opposite zeta potential rather than relying on the pore size of the membrane (see Figure 6.18).

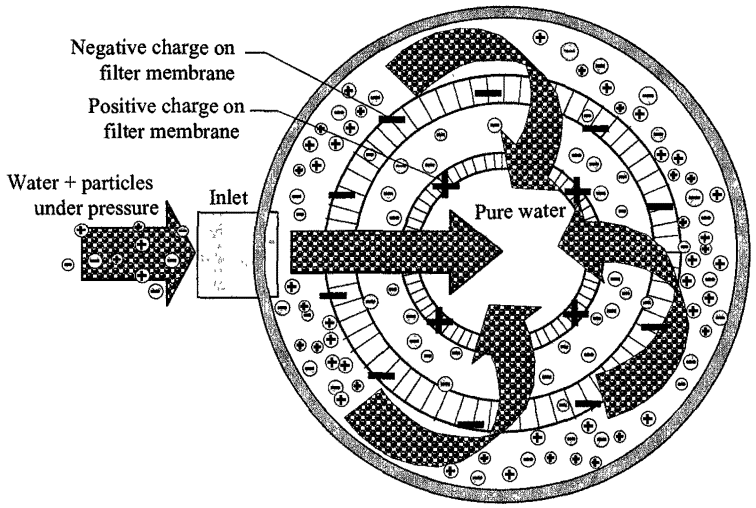


FIGURE 6.18 Electrostatic Filtration Using the Zeta Potential

Bacteria Control. Ultrapure water systems use *ultraviolet (UV) lamps* for bacterial sterilization. Water is exposed at a controlled rate to UV light waves, which deactivates certain molecules to reduce bacteria. UV systems are simple and reliable and can reduce bacteria to less than 1%.

Biological control of an ultrapure electronic-grade water system is also done by ozonating the water. Ozone (O_3) is created by discharging an electric current through dry air. The resulting ozone mixture is purified and then injected into the ultrapure water to kill the bacteria. All ozone is removed from the water by exposing the ozonated water to ultraviolet light, which breaks down the ozone to oxygen. Since most bacteria have a physical diameter in excess of 0.2 microns, point-of-use microfiltration can also be effective in removing bacteria.

Process Chemicals

Liquid chemicals used in semiconductor processing must be free of contamination for successful device yield and performance. Chemical purity is identified by an *assay number*, which describes the percentage of a particular chemical in the container (without reference to other substances present). For example, an assay of 99.99% for a bottle of hydrofluoric acid means the bottle contains 99.99% hydrofluoric acid and 0.01% other substances.

Filters are used to prevent chemical degradation during delivery or to maintain chemical purity during a recirculation step. Filters should be located at appropriate locations (e.g., near input to the gas controller) and as close to the process chamber as possible through point-of-use filtration. The different filter classes are:

Particle filtration: Depth-type filtration (see Figure 6.19) for particles from approximately 1.5 microns and larger.

Microfiltration: Membrane filtration of a liquid that removes particles in range of 0.1 to 1.5 microns.

Ultrafiltration: Pressure-driven membrane process that rejects large molecules from approximately 0.005 to 0.1 microns.

Reverse osmosis: Also called hyperfiltration. A pressure-driven solution process which transports liquid through a semipermeable membrane with the exclusion of particles and metal ions as small as approximately 0.005 microns.

A *membrane filter* uses a thin membrane of polymer or ceramic with small penetrating pores as the filter medium (see Figure 6.20). The size and distribution of the pores are controlled throughout the membrane. The membrane serves as a barrier that permits the passage of materials only up to a certain size or shape as defined by the pores. It is also called a surface filter because it removes particles from the air stream through interception at its surface. Membrane filters are used in reverse osmosis, microfiltration, and ultrafiltration. A membrane filter is often used as a point-of-use filter and is placed just before the process tool to provide final filtration.

A good filter will not create a significant pressure drop of the required flow, will not introduce secondary contamination, and will be compatible with the chemical. The *filter efficiency* is the percentage of particles of a specific size and above that are stopped in the filter. For liquid filters in a ULSI process, a typical efficiency rating for a membrane filter is 99.9999999% of 0.02 micron particles and above (referred to as *nine nines efficiency*).

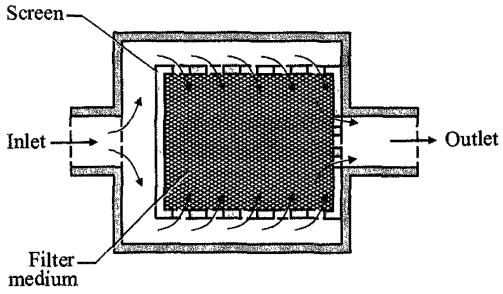


FIGURE 6.19 Depth-Type Filter (Used with permission from International SEMATECH)

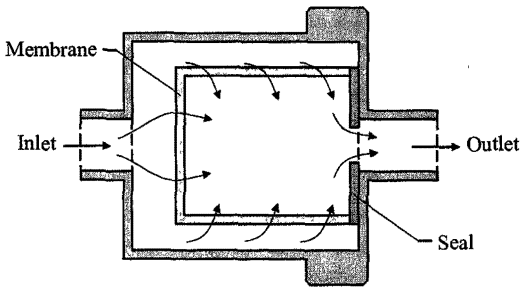


FIGURE 6.20 Membrane Filter (Used with permission from International SEMATECH)

Process Gases ■ The delivery and use of ultraclean gas is critical for semiconductor manufacturing in the ULSI era. However, the handling and delivery system can introduce impurities that adversely affect the yield of semiconductor devices. Gases pass through purifiers and gas filters to remove impurities and particles. Some gas filters are all-metal (e.g., nickel) that do not shed particles or outgas organic contaminants. These filters have a nickel membrane that is able to withstand the corrosive gases with a proven efficiency for particles as small as 0.003 microns. Other gas filters are made of a Teflon polymer.

Production Equipment

The production equipment used to manufacture the semiconductor wafer is the most significant source of particles in a wafer fab.[32] During the wafer fabrication process, silicon wafers are repeatedly loaded from cassettes into tools, processed through multiple equipment operations, unloaded back into cassettes, and then transported to the next workstation. This sequence occurs over and over for the 450 or more process steps necessary to fabricate a wafer, exposing the wafer to a multitude of mechanical and chemical operations at the different tools.

Many wafer fabrication operations occur in a vacuum, which requires special design considerations to avoid contamination. Examples of different sources of particle contamination from process tools are:

- Flaking of by-products built up on chamber walls.
- Automated wafer handling and transportation.
- Mechanical operations such as shaft rotation and valve opening or closing.
- Pumping and venting in vacuum environments.
- Cleaning and maintenance procedures.

Since equipment automation means fewer humans interacting with the product, there is less concern about particles from humans and more emphasis on reducing particles from equipment.[33] The number of particles on a wafer surface will increase during the fabrication process as the wafer is exposed to more tool operations. This condition is shown in Figure 6.21.[34]

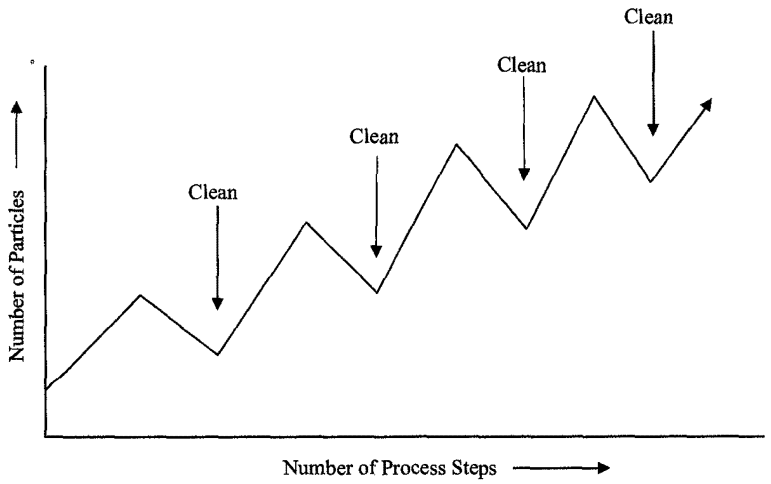


FIGURE 6.21 Wafer Surface Particles as a Function of Process Steps

Workstation Design

The right materials for *workstation design* of process tools and general workstations are necessary for achieving an ultraclean cleanroom. All materials emit some particles, but the goal is to reduce the emission to an acceptable level. Smooth, highly polished surfaces are best to reduce particulate contamination. Stainless steel has become a widely used material for work surfaces and equipment in the cleanroom. When properly processed, stainless steel has a relatively low particle emission rate. Electropolishing is a crucial finishing step.

Bulkhead Installation ■ Some fabs choose to install their equipment with a *bulkhead equipment layout.* In this approach, the major portion of the equipment is located behind the production bay fab wall in the service chase (see Figure 6.22 on page 132). Only the user-interface operator's controls and the wafer cassettes are located in the fab. This configuration isolates the equipment and its servicing in the chase, which is typically at a lower-class level of contamination.

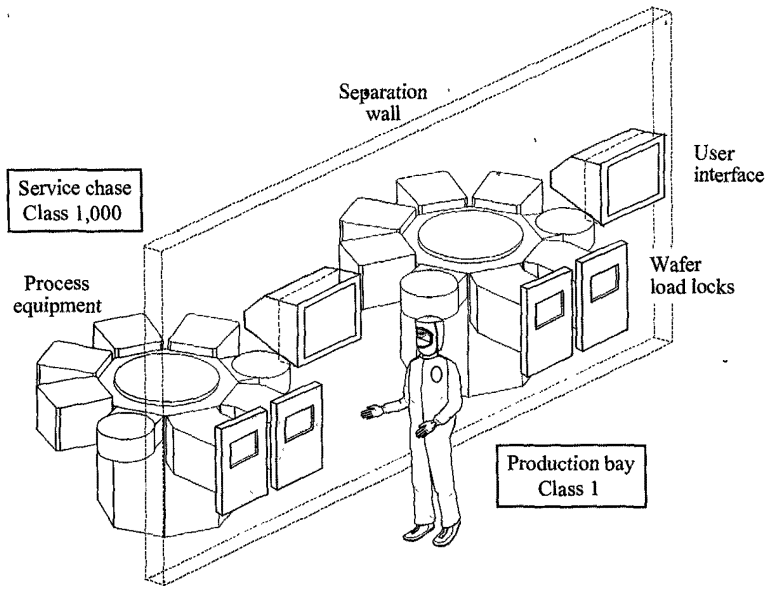


FIGURE 6.22 Bulkhead Installation

Handling ■ From the early days of semiconductor manufacturing until the 1970s, wafers were handled either manually with tweezers and vacuum wands (wafer holders that use vacuum to hold the wafer). Manual handling created particulate contamination and caused killer defects as device geometries decreased. Eventually manufacturers used *wafer cassettes* to transport wafers between tools (typically 25 wafers per cassette), with conveyor systems and elevators as a way to pick up and transfer wafers into and out of the tool (see Figure 6.23). Cassettes are designed to minimize particle generation, be electrostatic dissipative, and have minimal chemical outgassing.

Currently wafer handling within tool operations is largely done by robotic technology. *Robotic wafer handlers* do most cleanroom wafer handling to load and unload wafers from cassettes to the work area of the process tools and to manipulate wafers during the defined sequence of operations in a tool. This development has significantly reduced wafer particulate contamination over manual handling.

Once wafers arrive at a process chamber, they are placed on a chuck that holds the wafer during processing. In the early days of semiconductor manufacturing, this chuck was a mechanical clamp. It changed to a vacuum chuck to reduce particles on the topside of the wafer. However, a vacuum chuck tends to distort the shape of the wafer, which is undesirable during processing. In

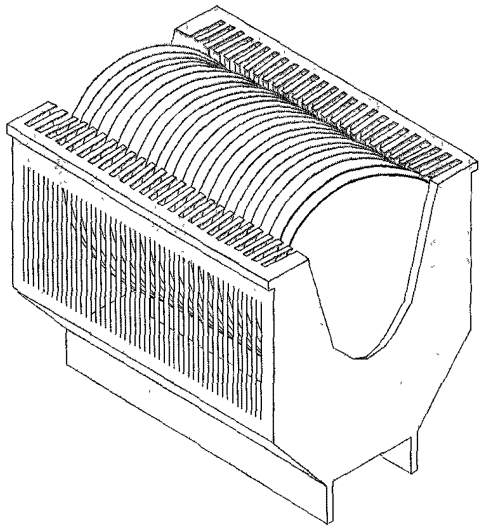


FIGURE 6.23 Wafer Cassette

order to improve process uniformity across the wafer surface, an *electrostatic chuck (ESC)* is commonly used today. It generates minimum particles and holds the wafer flat during processing (see Figure 6.24). An electrostatic chuck works by applying a voltage to an electrode on the chuck and generating a static charge. This electrode is isolated from the rear surface of the wafer by means of a dielectric material. An opposite charge is induced into the bottom side of the wafer, which pulls the wafer toward the chuck.

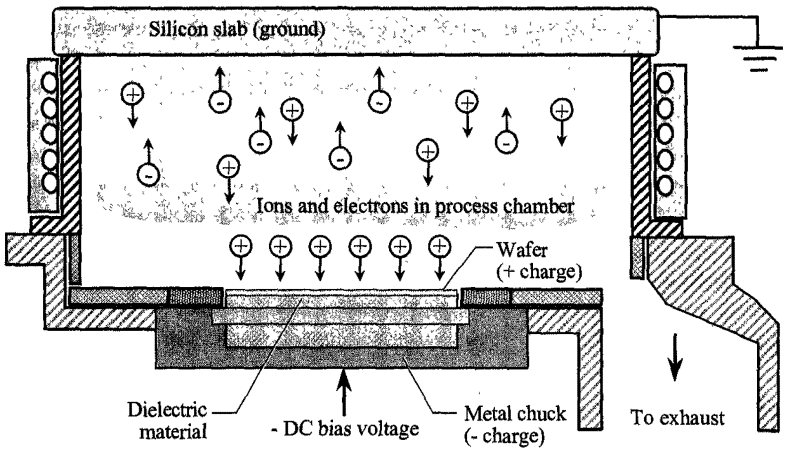


FIGURE 6.24 Electrostatic Wafer Chuck
(Used with permission from Applied Materials, Inc.)

Minienvironment ■ The concept of a cleanroom is continually being reassessed, primarily due to the need for more stringent control of contaminants and the high cost of constructing cleanrooms. There has been increased interest in controlling contamination at the specific workstation location where the wafer is processed through the use of minienvironments.

A *minienvironment* is a localized environment created by an enclosure that isolates wafers from the cleanroom environment while they are not in a process chamber (see Figure 6.25). This concept is also referred to as *wafer isolation technology.* A minienvironment clean zone can include the wafer cassettes used to hold wafers, wafer processing stations, loading ports, and storage locations.

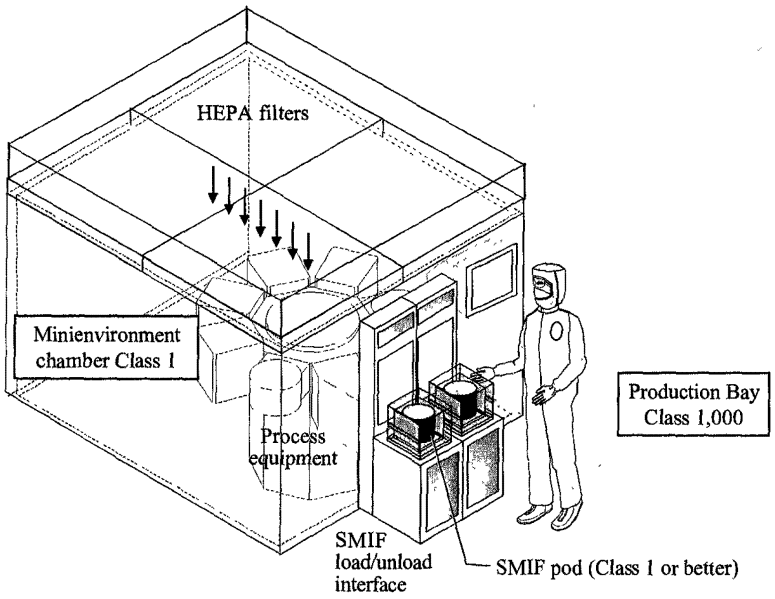


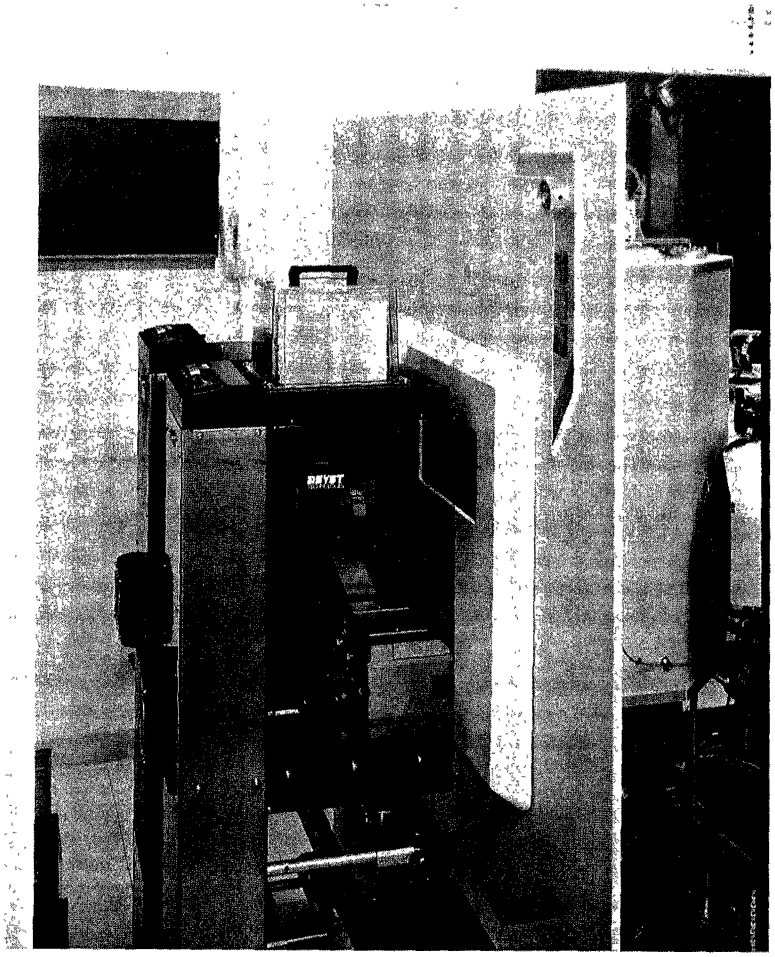
FIGURE 6.25 Minienvironment Concept

Minienvironments are controlled to an extremely pure cleanliness class (e.g., class 0.1 at 0.1 micron), while the cleanroom itself is at a higher class, such as a class 10. This condition makes it much simpler to achieve an ultraclean environment around the wafer during load/unload and processing at a workstation.

To transfer wafers between tools within a minienvironment, a standardized pod is used to enclose and transport cassettes of wafers. The pod has a *standard mechanical interface* (*SMIF*) to interface with the various tools, which was originally developed and patented by Hewlett-Packard. When a pod is presented to a tool, a robot on the tool automatically opens a door on the pod, removes the cassette, and presents it to the tool robot. SMIF systems can be added to existing tools or integrated into the tool enclosure.

For fabs processing 300-mm wafers, the cassette and pod will become one, permitting wafers to be handled directly from the pod. This new pod is referred to as a *front-opening unified pod* (*FOUP*). Handling will be completely automated with no manual lifting of wafer pods. The SMIF interface is fully integrated into the tool.

An advantage of a minienvironment is the possibility to control molecular contamination.[35] Wafers are isolated from molecular contamination through inert-gas purging of the minienvironment with nitrogen during loading, unloading, and transport. This control would be much more difficult to achieve in a large cleanroom. Minienvironments also reduce wafer exposure to water vapor during the nitrogen purge, which serves to inhibit native-oxide growth.



SMIF Pod on Bulkhead Installation
(Photo courtesy of Applied Materials, Inc.)

WAFER WET CLEANING

Due to the continual reduction in critical dimensions on a wafer, the wafer surface must be clean prior to undergoing processing. The most important way to control contamination on a wafer is to prevent contaminating the wafer. However, once a wafer surface is contaminated, then the contaminant must be removed by cleaning.

The goal for wafer cleaning is to remove all surface contaminants: particles, organics, metallics, and native oxides. Every wafer process step is a potential source of contamination to the devices on the wafer. Throughout the entire ULSI fabrication process, it is estimated that the surface of an individual wafer is wet cleaned up to 100 times.[36]

Wet-Cleaning Overview

The predominant wafer surface cleaning process is with *wet chemistry.* In the early 1980s, the consensus was that wet cleaning would be replaced by dry cleaning methods by the turn of the century. There has been a substantial effort to achieve this, but no completely successful replacement for wafer wet cleaning has been found. Wet cleaning of wafers is thriving and is being improved to attain more effective surface cleaning. The typical chemicals used in wafer wet cleaning and the contaminants they remove are shown in Table 6.5.[37]

TABLE 6.5 Wafer Wet-Cleaning Chemicals

Contaminant	Name	Chemical Mixture Description (all cleans are followed by a DI water rinse)	Chemicals
Particles	Piranha (SPM)	Sulfuric acid/hydrogen peroxide/DI water	$H_2SO_4/H_2O_2/H_2O$
	SC-1 (APM)	Ammonium hydroxide/hydrogen peroxide/DI water	$NH_4OH/H_2O_2/H_2O$
Organics	SC-1 (APM)	Ammonium hydroxide/hydrogen peroxide/DI water	$NH_4OH/H_2O_2/H_2O$
Metallics (not Cu)	SC-2 (HPM)	Hydrochloric acid/hydrogen peroxide/DI water	$HCl/H_2O_2/H_2O$
	Piranha (SPM)	Sulfuric acid/hydrogen peroxide/DI water	$H_2SO_4/H_2O_2/H_2O$
	DHF	Hydrofluoric acid/water solution (will not remove copper)	HF/H_2O
Native Oxides	DHF	Hydrofluoric acid/water solution (will not remove copper)	HF/H_2O
	BHF	Buffered hydrofluoric acid	$NH_4F/HF/H_2O$

One of the most critical surface cleaning process steps during wafer fabrication occurs prior to growing a thermal oxidation layer on the wafer. Ultrathin oxide layers must start with a completely clean wafer surface.

RCA Clean ■ The industry standard wet-clean process is referred to as the *RCA clean* process, developed by W. Kern and D. Puotinen at RCA in the 1960s and first published in 1970. RCA wet clean consists of sequential immersion in two different chemical baths, Standard Clean 1 (SC-1) and Standard Clean 2 (SC-2).

The chemical mixture of *Standard Clean 1 (SC-1)* is $NH_4OH/H_2O_2/H_2O$ (ammonium hydroxide/hydrogen peroxide/DI water). The three chemicals are mixed with a ratio range of 1:1:5 to 1:2:7. *Standard Clean 2 (SC-2)* is a composition of $HCl/H_2O_2/H_2O$ (hydrochloric acid/hydrogen peroxide/DI water), and is mixed in a ratio range of 1:1:6 to 1:2:8.[38] Both of these chemical mixtures are based on hydrogen peroxide (H_2O_2) and are traditionally used at a temperature of 75 to 85°C, with a 10 to 20 minute exposure time.

Standard Clean 1 (SC-1). As seen in Table 6.5, the SC-1 clean is an alkaline solution capable of removing particles and organic materials. For particles, the SC-1 wet-chemical process

works primarily through oxidation of the particle or by electric repulsion.[39] To understand the oxidation mechanism, realize that hydrogen peroxide is a powerful oxidizing agent that oxidizes the wafer surface and the particle. The oxidation layer on the particle can provide a liftoff mechanism that degrades and dissolves the particle, which breaks the adhesion forces between the particle and the surface. The particle then becomes soluble in the SC-1 solution and leaves the surface. This action is shown in Figure 6.26. This oxidizing action from the hydrogen peroxide also forms a protective layer on the silicon surface that keeps the particle from reattaching to the wafer surface.

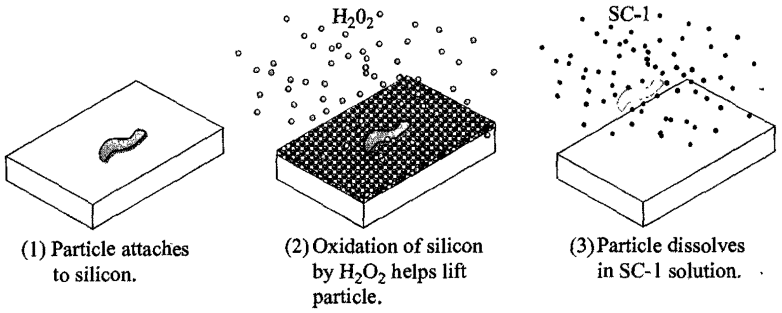


FIGURE 6.26 Oxidation and Solubility of Particle in SC-1

The SC-1 particle removal mechanism actually achieves an electrical repulsion of a particle. The hydroxide ion (OH^-) from ammonium hydroxide (NH_4OH) slightly etches the wafer surface and undercuts beneath the particle. The hydroxide ion also builds up a negative charge on the silicon surface and the particle. This negative charge on the particle and surface serves to repulse the particle from the surface and move it into the SC-1 solution (see Figure 6.27). Another benefit of the negative surface charge is that it prevents particles from redepositing.

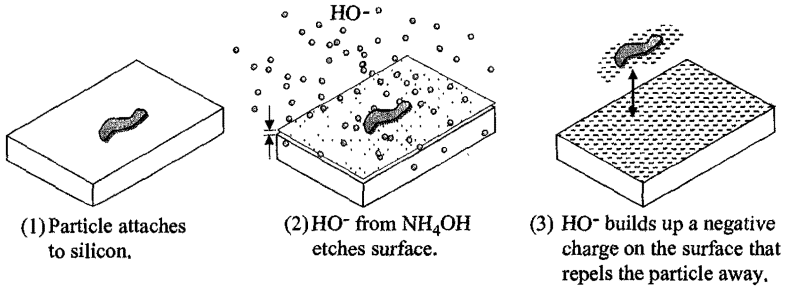


FIGURE 6.27 Particle Removal Through Negative-Charge Repulsion

Because the SC-1 step removes particles through an etching process on silicon, there is some microroughening of exposed silicon. This microroughening is a concern because it makes it difficult to grow very thin oxidation layers required in ULSI technology. Another concern regarding the use of SC-1 is a documented correlation between metal contamination on the wafer and the decomposition of the hydrogen peroxide in the mixture.[40] This contamination requires the bath to be replaced frequently, which is undesirable because it increases the use of chemicals in the fab.

It should be noted that the buildup of a charge potential on the wafer surface and the particle could be described by their zeta potential.[41] For particle removal, the zeta potential of the wafer and particle should be in the same polarity direction in the solution in order to have repulsive electrical forces between the wafer and particle.

Standard Clean 2 (SC-2). The SC-2 wet-clean process is used to remove metals from the surface of the wafer. To remove metallic (and some organic) contaminants from the silicon surface, it is necessary to have a solution with a high oxidation potential and low pH. In this manner, metals become ionized and dissolve in the acid solution that has a strong oxidizing action. Now the

cleaning solution can capture electrons from the metallic and organic contaminants and oxidize them. Metals are ionized to be dissolved in the solution while organic impurities are decomposed.

Modifications to RCA Clean ■ There have been modifications to the RCA clean process, mainly because this process uses chemicals and ultrapure water in high volume and at high temperature.[42] Very few chip manufacturers still use RCA clean in the same ratio as the original solution. Companies now will use mixtures that are up to 100 times more dilute with ultrapure H_2O, referred to as *dilute cleaning chemistries,* which achieves equal to or better cleaning effectiveness as the original solution.[43] For instance, a dilute SC-1 chemical ratio may be 1:4:50 for NH_4OH:H_2O_2:H_2O instead of the traditional 1:1:5 ratio. Dilute chemistries are an improvement for safety and health, plus they have the cost benefits of reduced chemical usage and disposal.

An important reason for the continued success of the RCA wet-clean is the availability of ultrapure water and chemicals. New cleaning approaches such as point-of-use chemical generation provide even higher levels of purity than ever before, which lead to more effective cleaning action. RCA clean does generate a large amount of chemical vapor, increasing the load on the cleanroom exhaust system to keep the chemical vapors from getting into the cleanroom. Another problem with bath evaporation is its effect on changing the bath composition over time.

Piranha Mixture. *Piranha* is a strong cleaning solution that combines sulfuric acid (H_2SO_4) and hydrogen peroxide (H_2O_2) to remove organic and metallic impurities on the wafer surface. Piranha is used at different steps in the process, sometimes before the SC-1 and SC-2 cleaning steps. The most common mixture is seven parts of concentrated H_2SO_4 to three parts of 30% (by volume) of H_2O_2. The usual cleaning method is to immerse the wafers in the piranha at 125°C for about 19 minutes followed by a thorough DI-water rinse. A variation of piranha is *Caro's acid,* which is prepared by mixing 380 parts of concentrated H_2SO_4 with 17 parts of 30% H_2O_2 and 1 part of ultrapure water.

HF Last Step. Many cleaning steps expose the wafer surface to a last step of hydrofluoric acid (HF) to remove native oxides on the wafer surface. Native oxide-free silicon is critical for producing a high-purity epitaxy film and ultrathin (50 angstroms and less) oxide layers for the gate region of MOS circuits. After the HF exposure, the wafer surface is completely terminated with hydrogen atoms and has a high stability against reoxidation in air.[44] A hydrogen-terminated silicon surface is maintained in the same condition as if it were bulk silicon crystal.[45]

Chemical Vapor Cleaning. Another method used in a few fabs is a chemical vapor to remove residual oxide and metallic contamination from a single wafer in a process chamber. The wafer is exposed to a fine mist spray of dilute HF:H_2O, followed by a DI-water rinse and an IPA (isopropyl alcohol) vapor-dry step. This method was developed to minimize HF chemical use, but it is not widely used because the increase in cleaning performance is minimal.

Wafer Cleaning Steps ■ A typical wafer cleaning sequence is shown in Table 6.6[46] on page 138. There are variations where some of the HF/H_2O steps are omitted.

TABLE 6.6 Typical Wafer Wet-Clean Sequence

Cleaning Step	Purpose
H_2SO_4/H_2O_2 (Piranha)	**Organics and Metals**
UPW rinse (ultrapure water)	Rinse
HF/H_2O (dilute HF)	Native oxides
UPW rinse	Rinse
$NH_4OH/H_2O_2/H_2O$ (SC-1)	**Particles**
UPW rinse	Rinse
HF/H_2O	Native oxides
UPW rinse	Rinse
$HCl/H_2O_2/H_2O$ (SC-2)	**Metals**
UPW rinse	Rinse
HF/H_2O	Native oxides
UPW rinse	Rinse
Drying	Dry

Wet-Clean Equipment

Because of the extensive use of wet cleaning in semiconductor manufacturing, the type of equipment used is a factor in reducing the chemical concentration and use of chemicals during wet cleaning. Traditional wet-clean processing has been done in *wet sinks,* consisting of a series of acid and rinse tanks housed in fume hoods. Self-contained cleaning equipment with microprocessor controls, robotic handling, and auto-dispensing of chemicals is common. The trend for wet cleaning and rinsing is for robotic handling of wafers with cassetteless operation within a minienvironment. Without the cassette, there is less obstruction of chemical flow to the wafer surface. This condition can improve the cleaning efficiency, reduce the amount of chemical usage, and lead to shorter rinse times.

Megasonics ■ One of the most widely used technologies with SC-1 for wet cleaning is megasonic cleaning. *Megasonic cleaning* uses ultrasonic energy with frequencies near 1 MHz during the clean process (see Figure 6.28). This process achieves much more effective particle removal at lower bath temperatures (30°C versus the original 80°C). This fact is important because of the difficulty in removing smaller particles, simply because of the difficulty in delivering the necessary force to such minuscule particle dimensions.[47]

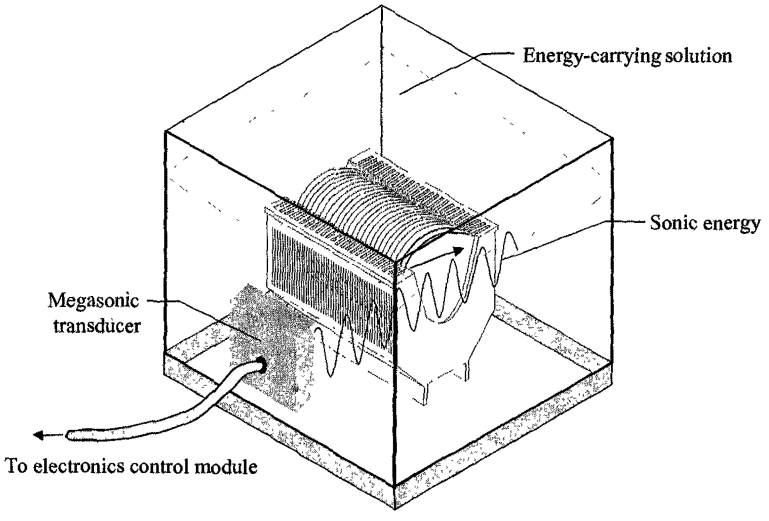


FIGURE 6.28 Megasonic Cleaning

Megasonic cleaning is accomplished when the vibrations of a megasonic transducer creates pressure waves in the liquid of the cleaning tank. The main particle removal mechanisms with megasonic cleaning are cavitation (formation of bubbles) and acoustic streaming.[48] *Cavitation* occurs when the low-pressure portion of the pressure waves create small bubbles that become filled with gas and/or vapor. These cavity bubbles oscillate through the liquid medium and violently collapse (implode) due to the sonic energy. This collapsing action is cavitation, which enhances particle removal without damage to the silicon wafer.[49] *Acoustic streaming* is the steady flow of liquid induced in the megasonic tank from the ultrasonic energy. Flowing liquid has more cleaning action than stationary water because it transports the particles away from the surface.

When the vibration frequencies are less than 100 kHz, then the process is referred to as ultrasonic. However, cavitation-induced pitting of silicon wafers occurs at ultrasonic frequencies that have not been found in the megasonic frequency range (800-1200 kHz).[50] Another reason for megasonics becoming more widespread in chemical cleaning and subsequent DI-water rinse operations is because it reduces the volume of chemicals required.

Spray Cleaning ■ With *spray cleaning* technology, the wet-cleaning chemicals are sprayed onto wafers, placed in a cassette and rotated inside a sealed chamber (see Figure 6.29). A DI-water rinse is sprayed on the wafers after each cleaning step and the DI water's resistivity is monitored to determine when all the chemicals are removed. The spray chamber is sealed during the process to isolate the chemicals and their vapors. After completion of the cleaning and rinsing cycles, the chamber is purged with hot nitrogen and the spin rate is accelerated to dry the wafers.

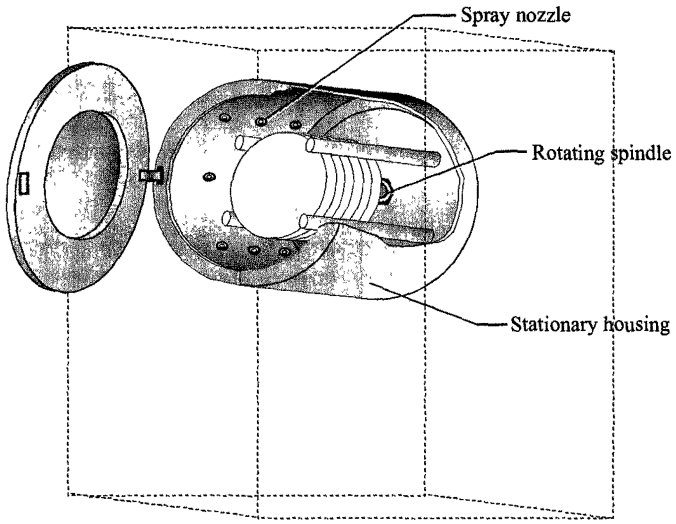


FIGURE 6.29 Spray Tool Designs for Wafer Cleaning

An advantage of spray tools is a continuous supply of premixed, freshly blended chemicals sprayed on the wafer. There is a physical force created during cleaning from the many small droplets of the spray hitting the wafer with a high turnover of chemicals and rinse water. The use of spray and centrifugal force from the rotating wafer ensures effective rinsing at a reduced chemical and water usage rate. There are also improvements in metallic and organic contamination removal when using spin-spray technology. However, spray cleaning does not provide uniformity of cleaning and rinsing because the center of the wafer is not turning at the high velocity of the outer wafer edge. This problem becomes worse with larger diameter wafers.

Scrubbers ■ Wafer *brush scrubbing* is an effective method for removing particles from the wafer surface. Brush scrubbing is widely used following wafer chemical mechanical planarization (known as CMP, which will be discussed in Chapter 18) due to the extensive particles generated during CMP. Brush scrubbing is able to remove particles one micron in diameter and smaller.

Early versions of wafer scrubbers made with nylon brushes damaged the wafer surface because of the stiffer nylon in conjunction with high pressure water sprays. Modern brushes are made with polyvinyl alcohol (PVA), which is a soft, compressible, spongelike material (see Figure 6.30). PVA brushes are effective at removing particles without wafer damage.[51] Brush scrubbers are available in a double-sided version that brushes both sides of the wafer at once. Brush scrubbers are often used with room temperature solutions of nontoxic chemicals or DI water that is sprayed on the wafer while brushing is underway.

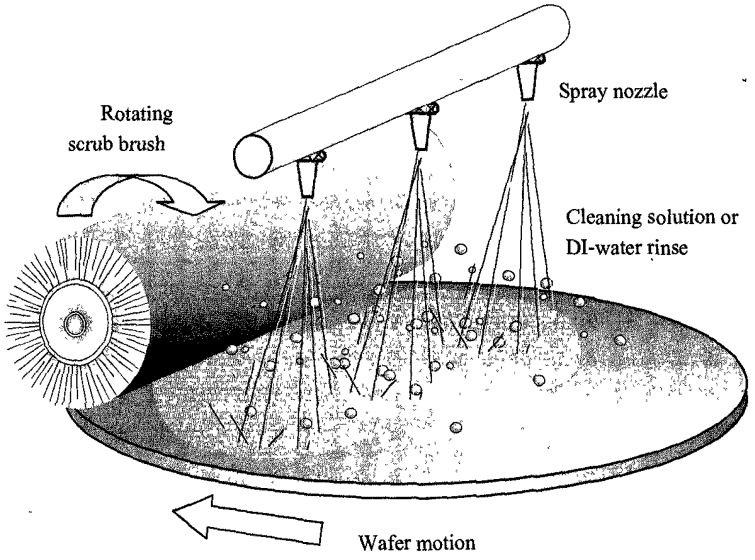


FIGURE 6.30 Wafer Brush Scrubber

Wafer Rinse ■ Chemical residues must be removed from the wafer after cleaning. Every wet-clean process step is followed by an ultrahigh purity *DI-water rinse.* This rinsing step requires a continuous supply of clean water to the wafer surface. The water rinse can also serve to end oxide-etching action from an HF cleaning step.

Overflow Rinsers. Traditionally the most common type of DI-water rinse has been the *overflow rinser* (see Figure 6.31). DI water is brought into the rinse system to flow through and around the wafers, sometimes with a nitrogen bubbler to aid in the mixing of chemicals at the wafer surface. The fluid motion in overflow rinsers serves to sweep away contaminants that have diffused into the flow stream from the surface of the wafers. High flow rates with no dead spots are goals for rinsing.

The overflow rinse process has also been applied to a cascade rinse system. In this case, the DI water cascades between two or three overflow rinsers connected to each other. The wafers start

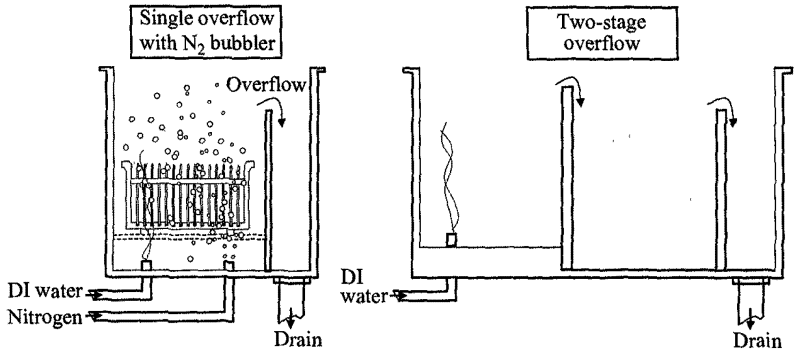


FIGURE 6.31 Overflow Rinser

the rinse process in the downstream rinser and are moved sequentially through the rinsers to the first one with the direct DI-water supply.

Overflow rinsers are the conventional system, but they consume a large amount of DI water. Given the large amount of water used in semiconductor facilities, there has been extensive investigation into alternative rinse processes that consume less water.

Dump Rinse. A common rinse method is the *dump rinser.* DI water is sprayed on the wafers while filling the rinse tank. At a certain water level, a drain in the bottom of the tank quickly opens and the water instantly dumps out (see Figure 6.32). The drain then closes and the cycle repeats for a set number of cycles. An inert gas, such as nitrogen, is often bubbled through the water to aid in removing contaminants such as particles by creating a scrubbing action.

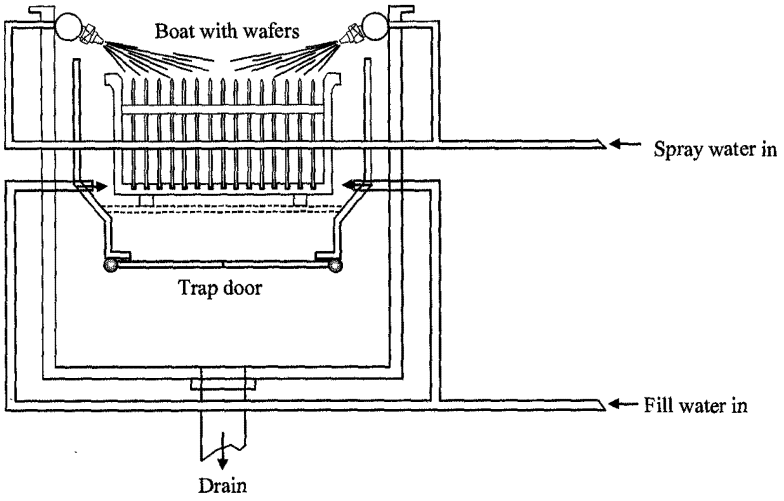


FIGURE 6.32 Dump Rinse

Spray Rinse. A *spray rinser* is typically used in conjunction with wafer cleaning and drying. It uses the physical force of the flowing water to dislodge residual chemicals on the wafer surface.

Hot DI-Water Rinsing. *Hot DI-water rinsing* (70 to 80°C) has become widely used for rinsing the wafer. The benefits of hot DI-water rinsing are that it aids in the removal of residual chemicals (especially if some type of HF solution has been used) and it improves the performance of the dry wafer. However, there has been some research that shows that the use of DI-water rinse at elevated temperatures creates a problem with etching the silicon surface, leading to surface microroughness.[52] Surface microroughness is undesirable for the extremely thin films required in ULSI technology.

Wafer Drying ■ Given the amount of aqueous rinse in a wafer process, it is important to dry the wafer, ideally with no drying spots. How a wafer responds to water refers to its wettability. Water will either adhere to a clean hydrophilic wafer or bead up on a hydrophobic surface. A *hydrophilic surface* has an affinity toward water, allowing water to spread across it in large puddles. Oxide-coated and RCA-cleaned wafer surfaces are hydrophilic. A *hydrophobic surface* has an aversion to water and does not support large pools of water. The water generally beads up, such as water on a newly waxed car, and is called *dewetted.* Oxide-free surfaces that have been hydrofluoric acid (HF) etched are hydrophobic due to the hydrogen-terminated surface. As the wafer undergoes processing, semiconductors require a hydrophobic surface to promote adhesion during subsequent layering processes.[53] Because HF-cleaned wafer surfaces are susceptible to contamination, there must be an adequate effort to properly dry the wafer surface.

Spin Dryers. The *spin dryer* has been widely used in the past. A cassette of wafers is placed in the spin equipment and high-speed rotation removes moisture while the wafers are being sprayed

with heated nitrogen. However, there are problems, such as the difficulty of removing moisture in holes and the generation of particulate contamination from the mechanism. In addition, the high-speed rotation of the wafer causes a charge buildup, which attracts particles. Adding static eliminators inside the equipment controls this static electricity buildup.

IPA Vapor Dry. In this process, wafer drying occurs through the displacement of water by heated solvent vapors of isopropyl alcohol (IPA). This *IPA vapor dry* method has a lower probability of contamination from particles, but the purity level of the IPA must be controlled. The IPA is heated in a tank with the wafers suspended in the vapor above the liquid level. When the wafers are removed from the vapor dryer, the solvent vapors evaporate, leaving the wafer dry.

Alternatives to RCA Clean

There are several alternative cleaning technologies to RCA clean that have been researched and used in varying degrees.

Dry Cleaning ■ Wet-cleaning methods are still meeting most cleaning requirements in wafer fabs. Dry-surface cleaning techniques are implemented primarily in integrated cluster tools, rather than stand-alone tools. Ultimately, the cost benefits through increased yield or reduced cost of ownership will decide the frequency of dry cleaning in wafer fabrication.

Plasma-Based Dry Cleaning. *Plasma-based dry cleaning* must be considered as a serious alternative to wet clean. Today, dry-plasma technology is used for removing organic photoresist from wafers (known as *ashing*), or as an integrated preconditioning step in process tools to remove native oxides. In the plasma process, gases and plasma energy are used to cause the chemical reaction that removes the contaminant (plasma processes will be discussed in detail in later chapters).

An example of plasma cleaning is a microwave downstream-plasma process.[54] This plasma can remove film residues containing organic material plus metallic and other inorganic components. The plasma can remove the sidewall film while minimizing removal of the underlying dielectric layer. This plasma clean is followed by a DI-water rinse.

There will be further development and use of dry-plasma cleaning processes for semiconductor manufacturing since the advanced processes required for 0.18-μm critical dimensions and below are using new materials with different forms of residues (within the same classes we have defined). In many cases, the chemical reactions of wet cleaning do not possess enough activation energy (energy required to cause a reaction) to remove all the residues. This is why the higher-activation energies associated with plasma techniques and nonplasma methods are needed for wafer cleaning.

Chelating Agents ■ A *chelating agent* binds and removes metallic ions. If this agent, such as ethylenediamine-tetra-acetic acid (EDTA) is added to a cleaning solution, it can reduce the redeposition of metals in the solution. It does this by altering the reduction-oxidation potential of the metallic species. An example would be adding a chelate to the ammonium hydroxide of SC-1 as a means of preventing metals from any chemical impurities from adhering to the wafer surface.

Ozone ■ *Ozone* (O_3) injected into ultrapure water has been identified as a possible replacement for some RCA wet cleaning. It has been shown that a combination of ozonized ultrapure water followed by an SC-2 cleaning step is effective in removing metals such as copper (Cu) and silver (Ag), as well as organic contaminants.[55] Research has indicated that ozone injected into DI water could replace a piranha step used for light organic cleaning.

Cryogenic Aerosol Cleaning ■ The principle of *cryogenic aerosol cleaning* is to sufficiently cool a gas (argon) to form solid ice particles that are injected onto a wafer surface to remove particulate contamination. The cooling of the argon is done by expansion cooling (lowering of pressure) in a vacuum chamber as a mixture of argon and nitrogen flows through a jet nozzle array. The nitrogen is used to dilute the argon and control the diameter of the solid argon particles. Research data has shown that cryogenic aerosol cleaning can be superior to wet cleaning. A benefit of this process is that it is environmentally benign. This process is not currently in wide use.

SUMMARY

Wafer fabrication is done in ultraclean cleanrooms. There are five types of contamination controlled in cleanrooms: particles, metallic impurities, organic contamination, native oxides, and electrostatic discharge (ESD), all of which can affect device performance. Particles must remain less than one-half the critical dimension or else they are known as a killer defect. Air is controlled by filtering and has a class number to designate the particle size and density. Humans must follow strict cleanroom protocol to minimize contamination. Facilities have special floor layouts to minimize the introduction of contaminants, with laminar airflow and HEPA filters used to attain ultraclean air. Air ionization is done to control ESD. Ultrapure DI water controls many types of contaminants through reverse osmosis, ultrafiltration, and bacteria control. Process chemicals and gases have different levels of filtering, transportation, and handling procedures to achieve high purity. Cleanroom equipment has special workstation designs to minimize contaminants and is becoming more controlled through use of the minienvironment.

The predominant method of wafer cleaning is a wet process using the RCA cleans of SC-1 and SC-2. Particles and organics are removed by SC-1 and metallics are removed by SC-2. Additional wet cleans are the piranha mixture and HF last step. Two common cleaning methods with RCA clean are megasonics and spray cleaning. Brush scrubbers are often used at the chemical mechanical planarization (CMP) operation to remove particles. Different types of DI-water rinse procedures are overflow rinsers, dump rinsers, spray rinsers, and hot DI water. Wafers are dried by spin dryers or IPA vapor dry. Alternatives to RCA wet clean are dry cleaning with plasma and the use of chelating agents, ozone, and cryogenic aerosol cleaning.

KEY TERMS

cleanroom
contamination
killer defects
particles
aerosols
particle density
particles per wafer per pass (PWP)
metal impurities
mobile ionic contaminants (MICs)
organic contaminants
native oxide
electrostatic discharge (ESD)
class number
ultrafine particles
cleanroom protocol
ballroom layout
bay and chase layout
sub-fab
laminar airflow
high-efficiency particulate air (HEPA)
ultra-low penetration air (ULPA)
air ionization
ultrapure DI-water
total organic carbon (TOC)
pyrogens
silica
makeup loop
polishing loop
water deionization
reverse osmosis (RO)
hyperfiltration
ultrafiltration
membrane contactor
zeta potential
ultraviolet (UV) lamps
assay number
filters
membrane filter
filter efficiency
workstation design
bulkhead equipment layout
wafer cassettes
robotic wafer handlers
electrostatic chuck (ESC)
minienvironment
standard mechanical interface (SMIF)
front-opening unified pod (FOUP)
wet chemistry
RCA clean
Standard clean 1 (SC-1)
Standard clean 2 (SC-2)
dilute cleaning chemistries

piranha
Caro's acid
wet sinks
megasonic cleaning
cavitation
acoustic streaming
spray cleaning
brush scrubbing
DI-water rinse
overflow rinser
dump rinser
spray rinser
hot DI-water rinsing
hydrophilic surface
hydrophobic surface
spin dryer
IPA vapor dry
plasma-based dry cleaning
chelating agent
ozone (O_3)
cryogenic aerosol cleaning

REVIEW QUESTIONS

1. Give a general description of a cleanroom.
2. What is contamination in semiconductor manufacturing?
3. Define a killer defect.
4. State the five categories of cleanroom contamination.
5. What is a particle? What is an aerosol? Why are particles a problem for semiconductor manufacturing?
6. Explain the rule of thumb for acceptable particle size in semiconductor manufacturing.
7. Describe the particle density on a wafer surface. What is PWP?
8. Give an example of a typical metal impurity. Give the two basic ways for metallics to deposit on a wafer surface.
9. Describe what an MIC is.
10. Identify a problem that results from metallic impurities in semiconductor manufacturing.
11. What is an organic contaminant? Provide two possible sources of organic contamination in wafer fabrication.
12. Identify two problems that result from organic contamination in semiconductor manufacturing.
13. Explain native oxide. Identify three problems that result from native oxide.
14. Explain electrostatic discharge (ESD).
15. Give three different problems in wafer fabrication resulting from ESD.
16. List seven sources of contamination in a wafer fabrication facility.
17. Explain the class number for air quality.
18. Interpret the following: (a) class 10 at 0.3 μm and (b) class 1 at 0.5 μm.
19. Describe an ultrafine particle.
20. Explain how a human can generate particles.
21. List four criteria that cleanroom garments should meet. Describe the modern cleanroom garment.
22. List seven proper cleanroom protocols that should be followed by cleanroom personnel.
23. Give three strategies for eliminating particles from cleanrooms.
24. Describe the ballroom layout for cleanrooms.
25. What is the bay and chase approach to cleanrooms?
26. What is laminar airflow, and at what class number does laminar airflow become critical?
27. How often may air turnover in a modern wafer fab?
28. What are HEPA and ULPA filters? What is the efficiency of an ULPA filter?
29. What are typical temperature and relative humidity (RH) settings in a wafer fab? Why is RH control important?
30. List and explain three ESD control methods.
31. Briefly describe each of the six unacceptable contaminants in ultrapure DI water.
32. What are the two major parts of a DI-water installation?
33. Explain the deionization of water. At what resistivity level is water considered deionized?
34. Describe reverse osmosis (RO) filtration. What is ultrafiltration?
35. Explain how a membrane contactor filter works.
36. Explain zeta potential.
37. How is bacteria controlled in ultrapure DI water?
38. What is the assay number for chemical purity?
39. List and discuss four classes of filters.
40. What is a membrane filter?
41. Describe the filter efficiency. What is the typical filter efficiency of a membrane filter?
42. Describe the filtration of a process gas.
43. What is the most significant source of particles in a wafer fab?
44. Give four examples of particle contamination from process tools.
45. What is a bulkhead equipment layout?
46. Describe how wafer cassettes reduced contamination over manual handling.
47. How is most cleanroom wafer handling done?
48. Explain the function and purpose of an electrostatic chuck.
49. Describe a minienvironment, and explain why this situation improves contamination control in a cleanroom.
50. What is a SMIF? What is a FOUP? How do they benefit wafer fabrication?

51. What is the goal for wafer cleaning?
52. What is the predominant wafer surface cleaning process?
53. Describe the RCA cleaning process.
54. What chemical mixture is used in SC-1? What contaminants are removed by SC-1?
55. Describe how the SC-1 wet-clean process removes wafer surface particles.
56. State two concerns that arise from cleaning with the SC-1 wet-clean process.
57. What chemical mixture is used in SC-2? What contaminants are removed by SC-2?
58. Explain what are dilute cleaning chemistries.
59. What is a piranha mixture, and what contaminants are removed from the wafer?
60. Discuss the HF last cleaning step and why it is used.
61. List a typical wafer wet-clean sequence. What is a wet sink?
62. Describe megasonics cleaning and why it is used, including cavitation and acoustic streaming.
63. Discuss spray-cleaning technology. What is an advantage from this cleaning method?
64. Describe wafer brush scrubbing. What contaminant does it remove, and at what process step is this method often used?
65. When is a DI-water rinse done? Describe three different methods for wafer rinsing.
66. Explain the difference between a hydrophilic and a hydrophobic surface.
67. Describe two different methods for wafer drying.
68. Discuss plasma cleaning. Why will the usage of dry plasma cleaning increase?
69. What is a chelating agent, and how is it used in cleaning?
70. How could ozone be used for wafer surface cleaning?
71. Describe cryogenic aerosol cleaning.

CHEMICAL AND EQUIPMENT SUPPLIERS' WEB SITES

Adept Technology Inc.	http://www.adept.com
AERONEX Inc.	http://www.aeronex.com
Air Kontrol Inc.	http://www.airkontrol.com/
Air Products and Chemicals	http://www.airproducts.com
Amerimade Technology	http://www.amerimade.com/
Apex Industries	http://www.apexind.com/
Applied Science and Technology	http://www.astex.com/
Aquionics Inc.	http://www.aquionix.com/
Asahi/America Inc.	http://www.asahi-america.com/
Ashland Specialty Co.	http://www.ashchem.com/
ASI Technologies	http://www.asidoors.com/
AST Products	http://www.astp.com
BOC Edwards	http://www.boc.com/edwards/
Clean Air Products	http://www.cleanairproducts.com/
Contamination Control Products	http://www.ccpcleanroom.com/
Dow Chemical/Filmtec	http://www.dow.com/liquidseps/
Dow Corning	http://www.dowcorning.com/
Dryden Engineering Co.	http://www.drydeneng.com/
Eastman Chemical Co.	http://www.eastman.com/
EKC Technology Inc.	http://www.ekctech.com/
Entegris Inc.	http://www.entegris.com/
Environflex Inc.	http://www.enviroflex.com/
Filtration Technology Inc.	http://www.filtrationtechnology.com/
FSI International	http://www.fsi-intl.com/
General Chemical Corp.	http://www.genchem.com/
IN USA Inc.	http://www.inusaozone.com
Integrated Designs LP	http://www.pumpless.com/
Ion Systems	http://www.ion.com/
J. T. Baker	http://www.jtbaker.com/
Kappler Protective Apparel	http://www.kappler.com/
Koch Microelectronics Service Co.	http://www.kochmicroelectronic.com/
Meissner Filtration Products Inc.	http://www.meissner.com/

MicroChem Corp.	http://www.microchem.com/
Micro Magazine	http://www.micromagazine.com/
Millipore Corp.	http://www.millipore.com/
Modutek Inc.	http://www.modutek.com
NetMotion Inc.	http://www.netmotion.com/
Pall Corp.	http://www.pall.com/
Parker Hannifin Corp.	http://www.parker.com/
Pope Scientific Inc.	http://www.popeinc.com/
PTI Advanced Filtration Inc.	http://www.pti-afi.com/
PURAC America Inc.	http://www.purac.com/
Schumacher	http://www.schumacher.com
SCP Global Technologies	http://www.scpglobal.com/
Sage Technologies Corp.	http://www.sagetech.net/
Semitool	http://www.semitool.com/
Simco Static Control	http://www.simco-static.com/
TEL, Tokyo Electron Ltd.	http://www.teainet.com
The Texwipe Co. LLC	http://www.texwipe.com/
Ultrapure & Industrial Services	http://www.ultrapure.com/
US Filter/Filterlite	http:///www2.usfilter.com/
Verteq Inc.	http://www.verteq.com/

REFERENCES

1. T. Hattori, "Particle Reduction in VLSI Manufacturing," *Contamination Control and Defect Reduction in Semiconductor Manufacturing III* (Pennington, NJ: The Electrochemical Society, 1994) p. 3.
2. A. French, *Newtonian Mechanics,* (New York: W. W. Norton, 1971), p. 32.
3. C. Gross et al., "Assessing Future Technology Requirements for Rapid Isolation and Sourcing of Faults," *Micromagazine* online edition (July/August 1998).
4. T. Hattori, "Detection and Analysis of Particles in Production Lines," *Ultraclean Surface Processing of Silicon Wafers: Secrets of VSLI Manufacturing,* ed. T. Hattori, (Berlin: Springer, 1998), p. 245.
5. W. Kern, *Handbook of Semiconductor Wafer Cleaning Technology,* (Park Ridge, NJ: Noyes Publications, 1993), p. 9.
6. C. Chang and T. Chao, "Wafer-Cleaning Technology," *USLI Technology,* ed. C. Chang and S. Sze (New York: McGraw-Hill, 1996), p. 67.
7. G. Cooper, "The Effect of CO Cylinder Materials on Wafer Contamination," *Semiconductor International* (July 1997): p. 301.
8. C. Chang and T. Chao, "Wafer-Cleaning Technology," p. 67.
9. S. De Gendt et al., "Silicon Surface Metal Contamination Measurements Using Grazing-Emission XRF Spectrometry," *Science and Technology of Semiconductor Surface Preparation, Symposium Proceedings* 477 (Warrendale, PA: Materials Research Society, 1997) p. 397.
10. K. Saga, "Influence of Surface Organic Contamination on the Incubation Time in Low-Pressure Chemical Vapor Deposition of Silicon Nitride on Silicon Substrates," *Science and Technology of Semiconductor Surface Preparation Symposium Proceedings* 477 (Warrendale, PA: Materials Research Society, 1997): p. 379.
11. T. Ohmi, "Total Room Temperature Wet Cleaning for Si Substrate Surface," *Journal of the Electrochemical Society* (September 1996): p. 2957.
12. C. Chang and T. Chao, "Wafer-Cleaning Technology," p. 80.
13. M. Suzuki et al., "Etching Characteristics During Cleaning of Silicon Surfaces by NF3-added Hydrogen and Water-Vapor Plasma Downstream Treatment," *Science and Technology of Semiconductor Surface Preparation, Symposium Proceedings* 477 (Warrendale, PA: Materials Research Society, 1997): p. 167.
14. D. Tolliver, *Handbook of Contamination Control in Microelectronics: Principles, Applications and Technology,* (Park Ridge, NJ: Noyes Publications, 1988), p. 175.
15. L. Levit and J. Menear, "Measuring and Quantifying Static Charge in Cleanrooms and Process Tools," *Solid State Technology* (February 1998): p. 85.
16. A. Braun, "Cleanroom Technologies Continue to Keep Contamination at Bay," *Semiconductor International* (March 1998): p. 58.
17. R. Kraft, "Proper Cleanroom Protocol," *Semiconductor International* (March 1998): p. 73.
18. Summarized from R. Kraft, "Proper Cleanroom Protocol," *Semiconductor International* (March 1998): p. 73.
19. R. Jarvis and L. Armentrout, "Full-Fab Surface Particle Detection Improves Yields," *Semiconductor International* (June 1997): p. 199.
20. J. Schroeder, "Automation-Centric Processing Bay Layout," *Semiconductor International* (June 1997): p. 209.

21. H. Tseng and R. Jansen, "Cleanroom Technology," *ULSI Technology,* ed. C. Chang and S. Sze (New York: McGraw-Hill, 1996), p. 14.
22. S. Middleman and A. Hochberg, *Process Engineering Analysis in Semiconductor Device Fabrication,* (New York: McGraw-Hill, 1993), p. 125.
23. H. Tseng and R. Jansen, "Cleanroom Technology," p. 28.
24. D. Mahoney, "The Construction of a Class 1 Cleanroom: A Case Study," *Semiconductor International* (October 1997): p. 141.
25. H. Inaba, "Effect of Electrostatic Charge on Particle Adhesion on Wafer Surfaces," *Ultraclean Surface Processing of Silicon Wafers: Secrets of VLSI Manufacturing,* ed. T. Hattori (Berlin, Springer, 1998): p. 143.
26. T. Roche, T. Peterson, and E. Hanson, "Water Use Efficiency in Immersion Wafer Rinsing," *Science and Technology of Semiconductor Surface Preparation, Symposium Proceedings* 477, (Warrendale, PA: Materials Research Society, 1997): p. 527.
27. R. Mohindra and W. Kern, "New Process for Producing Particle-Free Deionized Water," *Semiconductor International* (July 1997): p. 191.
28. M. Dax, "Membrane Contactor Technology Gives PPB Dissolved Oxygen in Water," *Semiconductor International* (December 1996): p. 36.
29. H. Tseng and R. Jansen, "Cleanroom Technology," *ULSI Technology,* ed. C. Chang and S. Sze (New York: McGraw-Hill, 1996), p. 39.
30. M. Dax, "Membrane Contactor," p. 36.
31. R. Mohindra and W. Kern, "New Process," p. 194.
32. T. Hattori, "Detection and Analysis of Particles in Production Lines," *Ultraclean Surface Processing of Silicon Wafers: Secrets of VLSI Manufacturing,* ed. T. Hattori (Berlin: Springer, 1998), p. 245.
33. A. Braun, "Cleanroom Technologies Continue to Keep Contamination at Bay," *Semiconductor International* (March 1998): p. 58.
34. T. Hattori, "Trends in Wafer Cleaning Technology," *Ultraclean Surface Processing of Silicon Wafers: Secrets of VLSI Manufacturing,* ed. T. Hattori (Berlin: Springer, 1998), p. 438.
35. R. McIllvaine, "Cleanroom Demands for the Next Generation," *Semiconductor International,* (March 1998): p. 68.
36. T. Ohmi, "Revolution of Silicon Substrate Surface Cleaning," *ULSI Science and Technology Proceedings* 97-3, (Pennington, NJ: The Electrochemical Society): p. 197.
37. T. Hattori, "Trends in Wafer Cleaning Technology," Supplement to *Solid State Technology* (May 1995): p. S8.
38. C. Chang and T. Chao, "Wafer-Cleaning Technology," p. 61.
39. Ibid., p. 64.
40. M. Meuris et al., "The IMEC Clean: A New Concept for Particle and Metal Removal on Si Surfaces," *Solid State Technology* (July 1995): p. 109.
41. T. Ohmi, "Total Room Temperature," p. 2957.
42. Ibid., p. 2958.
43. M. Heyns et al., "Advanced Wet and Dry Cleaning Coming Together for the Next Generation," *Solid State Technology* (March 1999): p. 40.
44. J. Rosamilia et al., "Hot Water Etching of Silicon Surfaces: New Insights of Mechanistic Understanding and Implications to Device Fabrication," *Science and Technology of Semiconductor Surface Preparation, Symposium Proceedings* 477 (Warrendale, PA: Materials Research Society, 1997): p. 181.
45. T. Ohmi, "Total Room Temperature," p. 2958.
46. T. Ohmi, "Total Room Temperature," p. 2961.
47. R. DeJule, "Trends in Wafer Cleaning," *Semiconductor International* (August 1998): p. 64.
48. J. Liu et al., "Si_3N_4 Particle Removal Efficiency Study," *Science and Technology of Semiconductor Surface Preparation, Symposium Proceedings* 477 (Warrendale, PA: Materials Research Society, 1997), p. 27.
49. G. Gale, A. Dai, and I. Kashkoush, "How to Accomplish Effective Megasonic Particle Removal," *Semiconductor International* (August 1996): p. 133.
50. Ibid.
51. R. DeJule, "Trends in Wafer Cleaning," p. 65.
52. J. Rosamilia, "Hot Water Etching," p. 181.
53. J. Park et al., "The Formation of Water Marks on Both Hydrophylic and Hydrophobic Wafers," *Science and Technology of Semiconductor Surface Preparation, Symposium Proceedings* 477 (Warrendale, PA: Materials Research Society, 1997): p. 513.
54. K. Lao and W. Wu, "Microwave Downstream Plasma Removes Metal Etch Residue," *Semiconductor International* (July 1997): p. 231.
55. T. Ohmi, "Total Room Temperature," p. 2959.

CHAPTER 7

METROLOGY AND DEFECT INSPECTION

Inspection has been done since the beginning of wafer fabrication. Semiconductor manufacturing veterans can recall when simply observing the color of the oxide film on the wafer surface was acceptable to estimate film thickness. Whichever color tint the oxide film appeared to be was compared to a color chart made up of pieces of wafers with different film thickness associated with each color tint (see Appendix D).

Wafer process inspection technology has undergone major changes. Feature sizes continue to shrink below the quarter micron era and chip density on wafers is continually increasing. Each process step has critical issues that define success or failure: contamination, junction depth, and film quality, to name a few. Furthermore, the introduction of new materials and processes will also bring new ways chips can fail. Measurements are critical to characterize the wafer materials and verify their acceptability.

To maintain good process yield and improve device performance, wafer fabs have improved control over process parameters and reduced sources of defects during fabrication. Some of these improvements have come from areas such as equipment automation, robotic handling, contamination reduction, and more consistent flow of wafers through the fab to avoid long waiting periods. Other improvements could not have occurred without the ability to measure wafers and evaluate the process performance. This evaluation is done using highly accurate measurement equipment that provides real-time data about the wafer fabrication performance, giving key information to the engineer and technician for the decision-making process.

OBJECTIVES

After studying the material in this chapter, you will be able to:

1. Explain why IC metrology is performed, and discuss equipment, yield and data collection issues associated with metrology.
2. Identify twelve different quality measures used in wafer fabrication, and identify the fabrication processes where each is used.
3. Describe the various metrology methods and equipment associated with the different quality measures.
4. List and discuss the purpose of seven different types of analytical equipment used to support IC fabrication.

INTRODUCTION

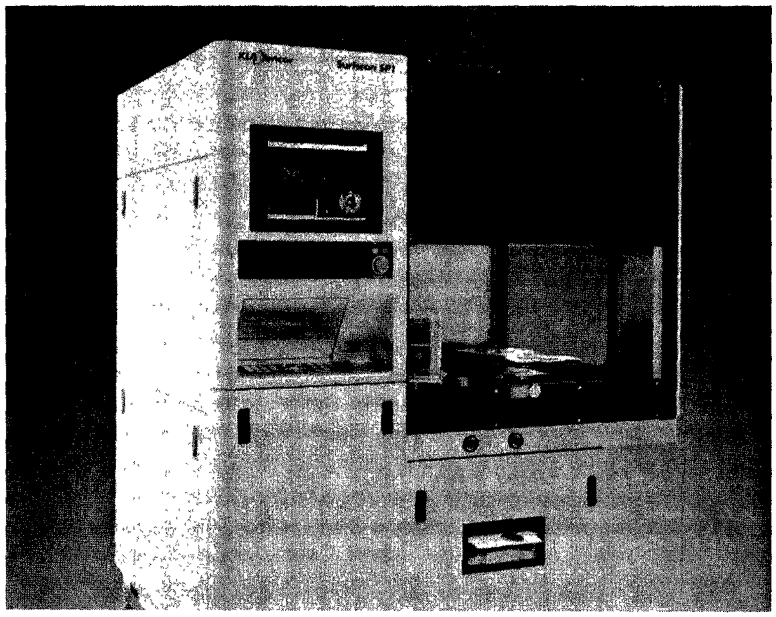
Metrology is the science of measurement to determine dimensions, quantity, or capacity. For IC fabrication, *metrology* refers to the techniques and procedures for determining physical and electrical properties of the wafer during the fabrication process. Metrology used during fabrication employs measurement equipment and sensors to collect and analyze data about wafer parameters and defects. *Defects* are the characteristics of the wafer or the results of the wafer fabrication process that cause nonconformance to the specified wafer requirement. The wafer *defect density* is the number of defects per unit area of wafer surface, usually cm^2. Wafer defects are specified by type and dimension. Metrology measurements are used by manufacturing personnel (e.g., technicians,

engineers, and managers) to ensure product performance and to make meaningful decisions regarding changes to improve processperformance.

It is essential that accurate evaluation of wafers be performed throughout the fabrication process to verify that the product meets the specified requirements. To achieve this, every process step in wafer fabrication has stringent quality measures that define the requirements needed at each step in order for the die to pass electrical test and meet reliability specifications during usage. Quality measures require extensive data collection on either test wafers or production wafers to demonstrate that the chip product's process requirements have been met.

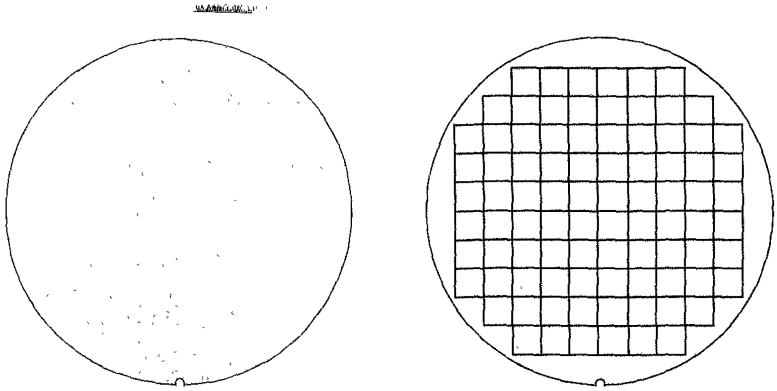
IC METROLOGY

IC metrology is a necessary means for measuring the performance of fabrication processes to ensure the specific quality standards are attained. To accomplish this requires a sample wafer, a measurement tool, and the means to analyze the data. Traditionally, most in-process data has been collected on *monitor wafers* (also referred to as *test wafers*) which are blank (or unpatterned) wafers included during process runs specifically to characterize the process. These blank wafers are recirculated for repeated usage (after undergoing appropriate refurbishing steps, such as surface stripping and cleaning). For instance, monitor wafers are included during thermal oxidation in a vertical diffusion furnace to measure oxide thickness and to check for particles inside the process chamber. Depending on the process step, blank wafers may include a predeposited film layer on the wafer surface.



Unpatterned Surface Inspection System
(Photo courtesy of KLA-Tencor Corporation)

Many semiconductor manufacturers have begun using production wafers, sometimes patterned wafers, for in-line tool monitoring (see Figure 7.1). Using actual production wafers more closely simulates what is happening during the process run, providing the manufacturing team members better information for making decisions.



Monitor wafer Patterned wafer

FIGURE 7.1 Monitor Wafer vs. Patterned Wafer

Measurement Equipment

There are different types of metrology equipment used to perform measurements in wafer fabrication. A major way to distinguish these tools is by how the measurement is done—with a stand-alone tool independent of the process or with metrology equipment integrated with the processing equipment. The two dominant categories for measurement equipment are shown in Table 7.1.[1] A *stand-alone measurement tool* functions independent of the process to perform metrology measurements. An *integrated measurement tool* has sensors that permit it to function as a part of the process and deliver in situ (real-time) measurement data.

Yield

Yield is an important indicator of a wafer fab's ability to produce high quality die. *Yield* is defined as the percent of good parts produced out of the total group of parts started. For example, if there are 200 die on a wafer and 190 die are acceptable, then the wafer yield is:

$$\text{Yield (\%)} = \frac{190}{200} \times 100 = 95\%$$

Different methods are used to calculate yield. One method is to measure yield relative to the type of part produced for a period of time (e.g., the yield of all wafers produced within a week). Yield can also be measured with respect to a particular location in the process, such as the die yield at the etch process. An important yield measurement for semiconductor manufacturing is wafer sort yield, which indicates the percentage of die that are acceptable after functional test (see Chapter 19).

Yield is widely used in semiconductor manufacturing to describe the health of the fabrication process. A high yield means that the process is producing good parts and performing as intended. A low yield is a strong indicator that there are quality problems in the product design or production process that must be resolved through improvement.

TABLE 7.1 Measurement Tool Classification

	Stand-Alone Tools
Off-line	Available only outside fab (usually destructive or contaminating).
At-line	Available in the fab (measures monitor wafers that are destructive, contaminated or unpatterned).
In-line	Used during production (can measure patterned wafers).
	Integrated Tools
On-line	Available at the process workstation to measure patterned wafers but not able to measure during wafer processing.
In situ	Measures wafer, process, or equipment during processing (real-time measurement).

In situ measurements are a recent development that were not very common even in the mid-1990s. As equipment efficiency improvements occur, there will be more integration of in situ metrology with fabrication equipment.

Data Management

Detecting defects is a challenge in IC fabrication because of shrinking feature sizes. The semiconductor manufacturer must have a method of categorizing the defects, sorting between false and real defects, and determining the root cause of each defect for determining the appropriate corrective action. Defect analysis should differentiate between random or nonrandom causes and correlate defect data with electrical and other tests to track how different defects affect wafer yield.

To a large degree, defect analysis is supported by an integrated software system. Large software programs connect major measurement tools and provide the ability to set up sample plans, analyze defect data, and detect trends amongst defects. Defects must be categorized in order to conduct failure analysis. The most advanced metrology software has *automatic defect classification* (*ADC*) to identify and classify defects based on software recognition patterns. Advanced defect management software also catalogs wafers and images based on user-defined criteria and includes statistical process control (SPC) capability for process monitoring and control.

QUALITY MEASURES

The breadth of metrology for IC fabrication is demonstrated by the many quality measures used throughout wafer processing. Semiconductor *quality measures* define the requirements for specific aspects of wafer fabrication to ensure acceptable device performance as measured by electrical tests and device reliability. Major quality measures used in the wafer fab are shown in Table 7.2, including the process areas where each measure is applicable. Semiconductor manufacturers specify their particular wafer quality measures by stipulating the exact requirements for their products at each process step. The manufacturer may specify other quality requirements beyond those listed in Table 7.2.

TABLE 7.2 Quality Measures in Wafer Fabrication by Production Areas

	Quality Measure	Implant	Diffusion	Thin Films		Polish	Etch	Photo
				Metals	Dielectric			
1	Film Thickness		√	√	√	√	√	√
2	Sheet Resistance	√	√	√				
3	Film Stress		√	√	√			
4	Refractive Index		√		√			
5	Dopant Concentration	√	√					
6	Unpatterned Surface Defects	√	√	√	√	√	√	
7	Patterned Surface Defects						√	√
8	Critical Dimensions (CDs)						√	√
9	Step Coverage					√	√	
10	Overlay Registration							√
11	Capacitance-Voltage		√					
12	Contact Angle							√

*The diffusion bay processes include: oxidation, deposition, diffusion, anneal, and alloy.

Film Thickness

Since wafer fabrication is a layering process, there are many different types of films found on the wafer surface throughout the fabrication process. Some of these different types of films are metal, dielectric, photoresist, and polysilicon. The quality of these films is essential to a high-yield fabrication process to produce reliable chips.

A critical quality parameter for films are their thickness. *Film-thickness metrologies* can be divided into two general types: whether they measure opaque (i.e., light-blocking, such as metal) films or transparent films. In some instances, such as for the gate oxide dielectric, the film thickness must be measured with an accuracy of an angstrom (Å) or less. Additional film quality parameters are surface roughness, reflectivity, density, and lack of pinholes and voids.

Resistivity and Sheet Resistance ■ One of the most practical methods for evaluating thickness for conductive films is to measure the *sheet resistance,* R_s. To discuss R_s, we need to understand how resistivity is related to resistance in a thin film layer. Consider current flow through a square sheet of conductive material, with the thickness, length, and width shown in Figure 7.2. Recall from Chapter 2 that resistance, R, of a conductor is given as:

$$R = \frac{\rho(l)}{a} \quad \text{(ohms)}$$

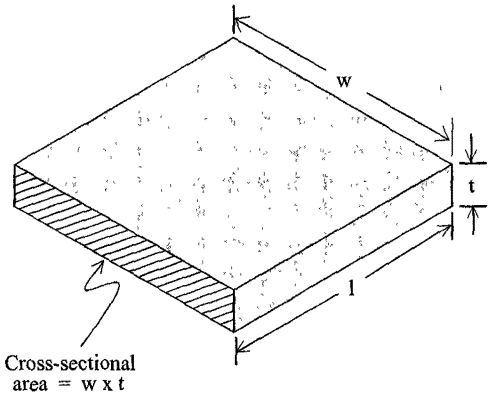


FIGURE 7.2 Illustration of Square Thin Film

The cross-sectional area is given by the product of width and thickness (w × t) and can be substituted in the equation for the area. Based on the assumption that the length (l) is the same dimension as width (w), then *l* can be substituted for *w*. Making these substitutions yields:

$$R_s = \frac{\rho(l)}{w \times t}$$

$$R_s = \frac{\rho(l)}{w \times t}$$

$$R_s = \frac{\rho}{t} \quad \text{(ohms/square, or } \Omega/\square\text{)}$$

Where,
R_s = Sheet resistance in ohms/square
ρ = Film resistivity in ohm-cm
t = Thickness of film

Sheet resistance, R_s, can be interpreted as the end-to-end resistance of a square sample of a thin film on a wafer. It depends on the film resistivity and thickness. Sheet resistance is independent of the size of the square film sheet. Measuring the sheet resistance between two points at the same distance

apart on any size square of the same thickness will produce the same resistance. For this reason, the units for R_s are given as ohms/square (Ω/□). Dimensionally the units for R_s are the same as ohms, but ohms/square is typically used as a reminder that this is actually the resistance of a sheet of film. If the thickness and sheet resistance are known, then the *sheet resistivity,* ρ_s, can be calculated:

$$\rho_s = R_s\,(t) \qquad \text{(ohms-cm)}$$

The terms sheet resistance and sheet resistivity are sometimes used interchangeably since they only differ by the thickness of the film. A tool used for measuring sheet resistance is the four-point probe.

Four-Point Probe. Resistance cannot be practically measured on thin films with two simple probes, as usually done with multimeters commonly used by electronic technicians. This is because of excessive contact resistance at the contact interface between the probes and the wafer material. In the semiconductor industry, a widely used method to measure sheet resistance is the *four-point probe.* This method has four in-line probes equally spaced that touch the wafer surface in a single file (see Figure 7.3). A known value of current (I) is passed between the two outer probes and the potential difference (v) developed across the two inner probes is measured. This approach avoids dealing with the effect of contact resistance.

The spacing between probes, s, should be less than the wafer diameter and less than the film thickness. The resistivity of the film sheet is related to the four-point probe current and voltage by:

$$\rho_s = \frac{V}{I} \times 2\pi s \qquad \text{(ohms-cm)}$$

Where,
ρ_s = Sheet resistivity in ohm-cm
V = DC voltage across the voltage probes (in volts)
I = Constant DC current passing through the current probes (in amperes)
S = Spacing between the probes

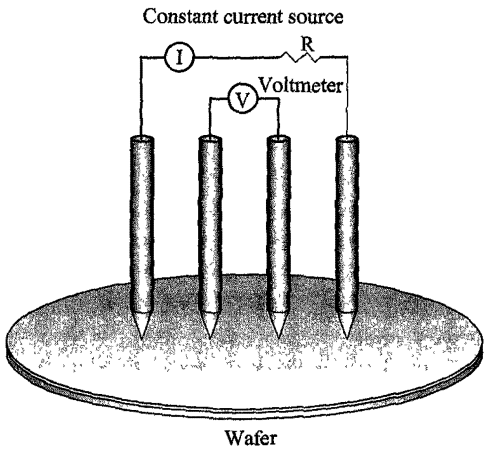


FIGURE 7.3 Four-Point Probe

Resistivity measurements with a four-point probe can be reproduced to within ± 2% if care is taken in selecting instrumentation, and in determining probe pressure and current levels.[2] The four-point probe is a nondestructive technique (e.g., the wafer does not have to undergo permanent damage to perform the test). However, due to the potential for damage to the wafer from the probe contacts, the semiconductor industry has moved toward noncontacting probes.[3] One contactless method is the use of eddy-current probes. A high-frequency ac current flowing through a coil induces eddy currents in conducting films placed under the coil. The eddy currents are a loss in energy that can be attributed to the loading effects produced by the resistance of the conducting film. The resulting change in electrical energy can be measured and used in computing the value of the sheet resistance of the film being measured.

Sheet Resistance (Opaque Films) ■ Sheet resistance is used to indirectly measure the thickness of opaque conductive films (e.g., metal, silicide, or semiconductor films) deposited on an insulator substrate such as a silicon wafer. As long as the film layer is large and the probe spacing is small, the sheet resistance, R_s, is given by:[4]

$$R_s = 4.53 \frac{V}{I} \quad \text{(ohms/square, or } \Omega/\square\text{)}$$

This formula is actually derived from the equation for sheet resistivity listed on page 154. The constant 4.53 results from a correction factor based on the assumption of an infinitely large sheet of film with small probe spacing. The correction factor can be modified to a different value if the film sheet is not infinitely large.

Van der Pauw. A modification of the four-point probe is the *Van der Pauw Method* that measures sheet resistivity with four probes on the periphery of an arbitrarily shaped sample (see Figure 7.4). The concept is based on measuring electrical current on two corners of a square pattern and measuring voltage on the other two. The results are the same as those achieved with the four-point probe.

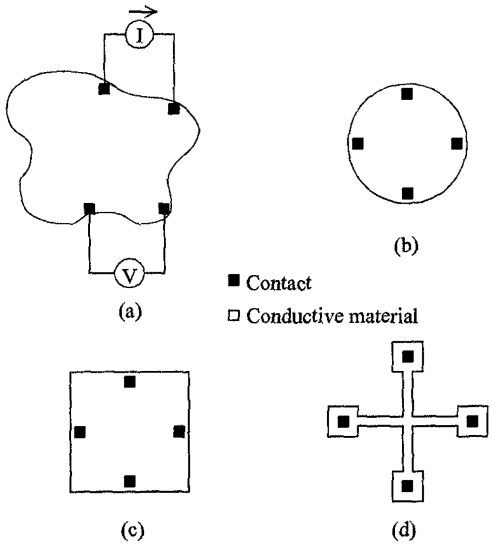


FIGURE 7.4 Van der Pauw Sheet Resistivity

Contour Maps. A practical benefit of sheet resistance measurements with the four-point probe is that the technique can be done at many sites across a test wafer to provide *contour maps* (see Figure 7.5). A contour map will typically present the data with a contour line showing the nominal $\Omega/\square$ and then deviations above and below this nominal value. The sheet resistance data results from a predetermined number of measurement sites across the wafer.

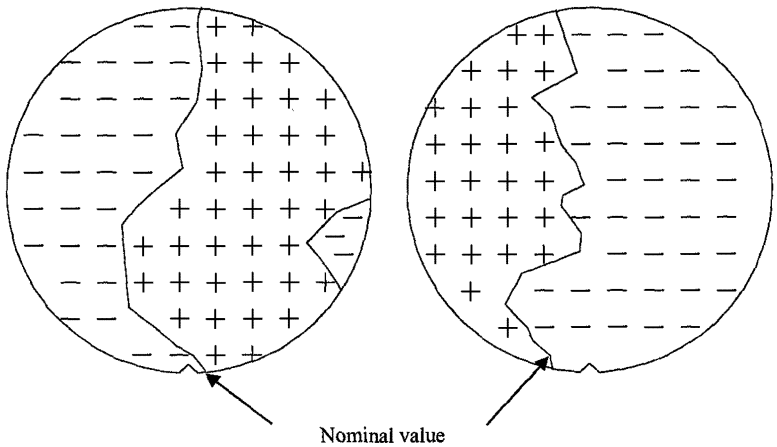


FIGURE 7.5 Sheet Resistance Contour Map

Ellipsometry (Transparent Films) ■ *Ellipsometry* is a nondestructive, noncontact optical film-thickness measurement technique primarily used to measure thin, transparent films. It is the leading method for film thickness metrology in the wafer fab.[5] The basic principle of ellipsometry is to use a linearly polarized laser light source that, when reflected from the sample, becomes elliptically polarized (see Figure 7.6). Polarized light consists of all light rays traveling in one plane. Ellipsometry measures the shape of this reflected ellipse, and, based on known inputs such as the angle of reflection, it accurately determines the film thickness. It is common in ellipsometry tools for the angle of the incident light to be varied to provide more measurement samples that are optimized for the film material. This practice is referred to as *variable-angle spectroscopic ellipsometry (VASE)*. VASE has improved ellipsometry for measuring multilayer stack structures, which is common in ULSI (see Chapter 12 for a description of multilayer stacks).[6]

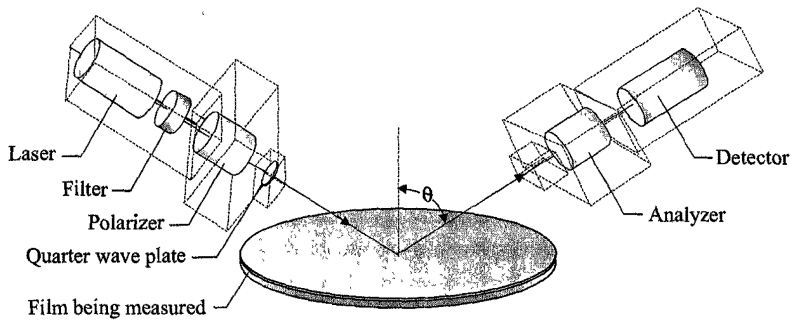


FIGURE 7.6 Basic Principle of Ellipsometry

Spectroscopic ellipsometry analyzes the polarization change of light over a wide spectral range (spectroscopy uses the characteristic light wavelengths emitted by different materials). Information about how the different wavelengths vary relative to one another after being reflected from the film is very useful in a production environment because it makes ellipsometry sensitive to different physical properties of the thin film. Ellipsometry is capable of measuring different types of very thin films on the order of tens of angstroms thick.[7] The types of films include dielectrics, metals, and polymer coatings. The main requirement is that the film be transparent or semitransparent. Fully automated thin film ellipsometry tools routinely measure gate oxide films < 40 Å thick with a repeatability better than 0.1%. Another example of ellipsometry measurements is sub-100 Å stacked oxide/nitride/oxide (ONO) film structures used in MOS capacitors with measurement repeatability in the 1% range. Thin metal layers (< 500 Å) are considered semitransparent and can be measured with ellipsometry. An example of a semitransparent metal layer is the thin copper seed layer used in copper interconnects. Metal layers (> 1000 Å) are generally considered opaque and cannot be measured with ellipsometry.

Ellipsometry measurement-tool features such as small measurement spots, pattern-recognition software, and high-precision wafer positioning hardware have made it possible to do many transparent film process-control measurements directly on production wafers. This reduction in unpatterned monitor wafers has reduced process-control costs and increased ellipsometry efficiency.[8] Ellipsometer tools are also being directly integrated into process tools for in situ (real time) measurement applications in areas such as etch and planarization. Real-time thickness measurement is desirable because it permits the process to define a precise endpoint of the film thickness during runs for better accuracy and repeatability in film thickness.

Reflection Spectroscopy ■ When light reflects off a surface, *reflection spectroscopy* (also referred to as *reflectometry*) is one of the three general types of optical measurement techniques, with the other two being optical microscopy (discussed later in this chapter) and ellipsometry (see above). Reflection of a structure is often used to characterize layer thickness for light-absorbing dielectric layers on a nonabsorbing wafer substrate (see Figure 7.7). Based on the relationship of how light reflects off the top and bottom surface of the film layer, reflectometry can be used to compute the film thickness. An advanced reflection-spectroscopy technique utilizes a dual-beam

spectrometer in order to give a clearer light signal for imaging. One light beam is used for reflection measurements from the film layer, while the second light source provides a reference source to correct for real-time lamp deviation or noise.

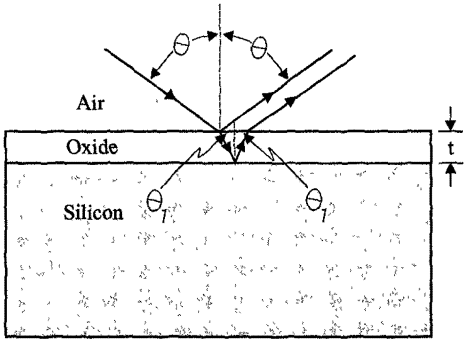


FIGURE 7.7 Light Reflection from a Thin Layer

X-Ray Film Thickness ■ X-ray beams can be focused on a surface and used for making film thickness measurements through a lesser-used technique known as *X-ray fluorescence (XRF)*. When X-rays strike a film, the absorption of radiation excites electrons in the film. The excited electrons drop into a lower energy state, emitting X-ray photons (known as fluorescence) whose energy represents the identity of the film atom. By measuring these X-ray photons, film thickness can be determined (see Figure 7.8). A modification is to use *total-reflection XRF (TRXRF)*, which uses a small angle and reduces the amount of X-ray scattering to improve the measurement sensitivity.[9]

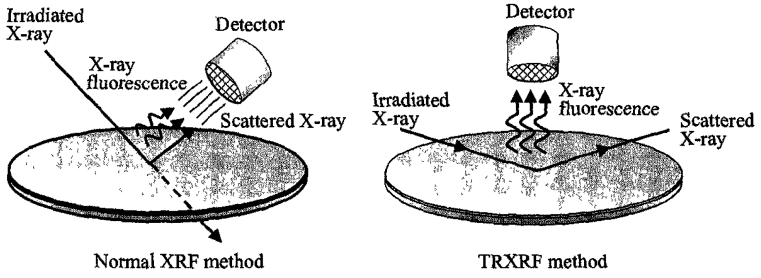


FIGURE 7.8 Film Thickness Measured with X-Ray Fluorescence (XRF)

A thin film's X-ray reflectivity will vary depending on the angle of incidence and wavelength. The reflectivity can be used to calculate the thickness and density of the film layer. It is difficult to use X-rays to measure films with mixed compositions or film stacks with two or more layers. X-ray fluorescence applies primarily to single-layer films.

Photoacoustic Technology ■ A recent development in metal thin film measurement is *photoacoustics*. This noncontact technology is based on light-induced sound pulses that generate an acoustic pulse that is directed toward the film stack. When the acoustic pulse strikes the surface and underlying film interface, an echo is created that bounces back toward the surface. This echo causes a slight change in reflectivity that is detected at the wafer surface (see Figure 7.9 on page 158). The time it takes for the pulse echoes to bounce back is used to calculate the film thickness.

This technique has a spot size of < 8 μm, which because of its small size, makes it capable of probing structures on patterned wafers. It can measure film stacks with an individual layer thickness down to < 20 Å, which is critical as device scaling requires smaller structures for increased performance.[10]

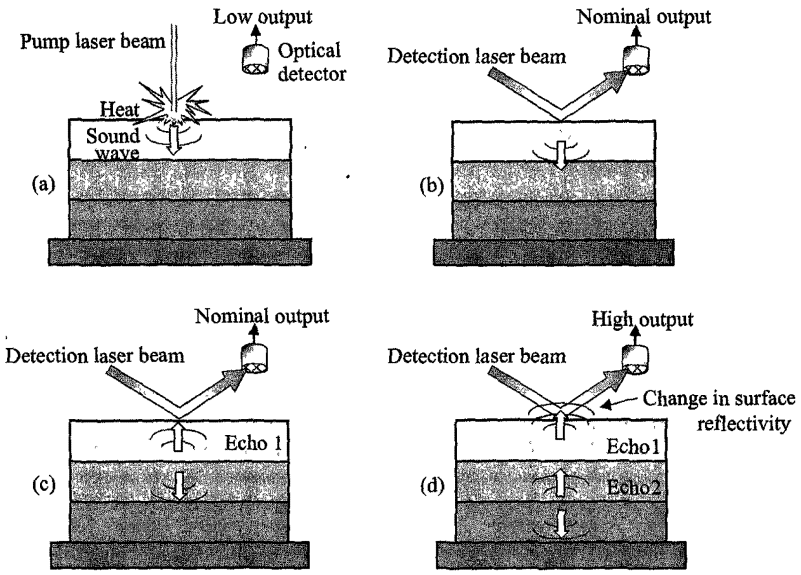


FIGURE 7.9 Photoacoustic Film-Thickness Measurement
Redrawn from *Solid State Technology,* (June 1997), p. 86.

Film Stress

Highly localized *film stress* can be introduced during routine manufacturing processes into thin films that may deform the substrate and create reliability concerns. The amount of this deformation is measured with thin film stress measurement tools. Stress measurement is done by analyzing the changes in the radius of curvature of the substrate resulting from the film deposition and applies to all types of standard thin films, including metals, dielectrics, and polymers. The wafer radius is measured before and after film deposition using either a scanning laser-beam technique or split-beam laser technique to create a stress profile map across the wafer (see Figure 7.10). Automated stress measurement tools have SMIF handling capability.

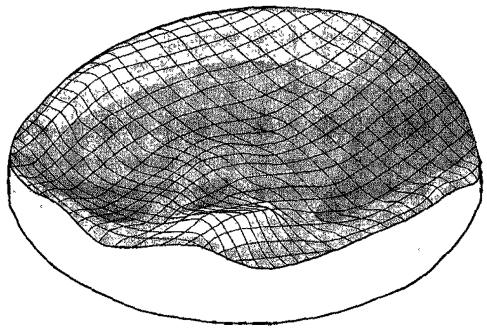


FIGURE 7.10 Detailed Stress Map of Wafer

Refractive Index

Refraction is the property of a transparent substance that addresses how much a light beam bends as it travels through it (see Figure 7.11). Variations in the refractive index can indicate contamination in the film and lead to incorrect thickness measurements. The index of refraction for pure oxide is 1.46. The index of refraction for thin films is measured by interference and ellipsometry techniques with the same ellipsometer measurement tool used to determine thin film thickness.

- Index of Refraction, $n = \frac{\sin \theta_i}{\sin \theta_r}$
- Examples of n:
  - air = 1.00
  - SiO_2 = 1.46
  - diamond = 2.12

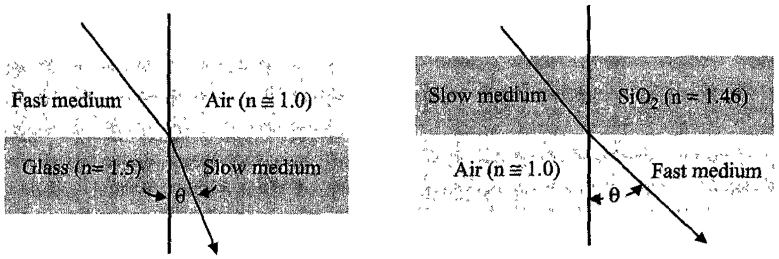


FIGURE 7.11 Index of Refraction

Dopant Concentration

The performance of semiconductor devices is directly affected by distribution of dopant atoms in the silicon in areas such as the formation of pn junctions, the epilayer, and the doping of polysilicon (see Figure 7.12). Modern processes employ dopant concentrations ranging from about 10^{10} atoms/cm2 to about 10^{18} atoms/cm^2 (see Chapter 17).[11] There are several techniques for measuring the dopant concentration, or dose, of atoms in silicon. A common in-line method is the four-point probe measurement, typically used for high-dopant concentration. The thermal-wave system is also used in-line and is acceptable for low-dose readings. For off-line measurements, secondary-ion mass spectrometry (SIMS, described later in this chapter) with whole-wafer positioning has recently been used as an alternative method for doping-concentration process control.[12] The capacitance-voltage test (described later in this chapter) can also be used to characterize dopant concentration.

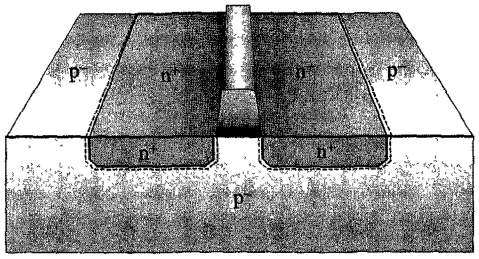


FIGURE 7.12 pn Junction

The four-point probe method has been previously discussed in this chapter, and is used in the same manner for dopant concentration as for sheet resistance (since sheet resistivity of the implanted silicon layer is related to the dopant concentration). The contour mapping technique possible with the four-point probe method is useful for routinely monitoring the implantation dose concentration.

Thermal-Wave System ■ A widely used method for monitoring ion-implant dose concentration is the *thermal-wave system.* This method measures the lattice damage in the implanted wafer due to ion implantation.[13] This is done by measuring changes in the wafer surface reflectivity from two lasers focused on the same localized spot on the wafer (see Figure 7.13 on page 160). One laser heats the wafer with a modulated Ar laser to create waves of heat (thus the term thermal wave). The thermal waves cause a change in the reflectance of the other HeNe probe laser that is proportional to the number of crystal defect sites in the wafer. A thermal-wave signal detector with calibration data is used to characterize the implant process by correlating the amount of crystal damage to the dopant concentration and other implant parameters.

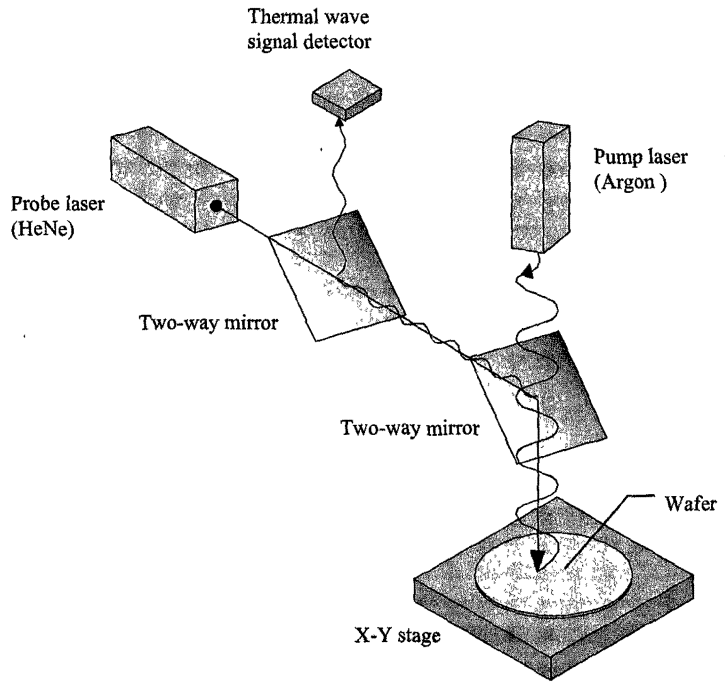


FIGURE 7.13 Thermal-Wave System for Measuring Dopant Concentration

The thermal-wave system has become common for ion-implant monitoring because of its ease of use and ability to be used on both patterned and unpatterned wafers. Its main disadvantage is that it measures damage and therefore needs calibration curves to indirectly assess dopant concentration.

Spreading Resistance Probe ■ The *spreading resistance probe* (*SRP*) has been a metrology tool in wafer fabrication since the 1960s and is used to measure both dopant concentration depth profiles and resistivity.[14] It is capable of profiling very shallow pn junction depths. The spreading resistance probe has two carefully aligned probes that are moved in steps along a beveled wafer surface, with the resistance between the probes measured at each step (see Figure 7.14). As the probes pass through the junction, the probes sense the change in conductivity type (n or p). The sample must be carefully prepared with a bevel angle, usually < 1°, which makes the SRP a destructive test. As the probes step through the wafer bevel, the spreading resistance, R_{sp}, of a flat circular contact of radius r on a planar surface of semiinfinite material of resistivity ρ is given by,

$$R_{sp} = \frac{\rho}{4r} \quad \text{(ohms)}$$

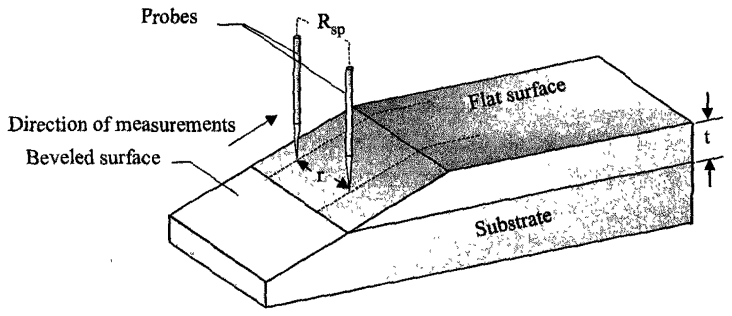


FIGURE 7.14 Spreading Resistance Probe (SRP)

The SRP employs calibration curves specific for a manufacturer's instrument to overcome a number of error sources associated with this measurement (e.g., silicon surface finish, contact resistance, tip deformation, bevel surface preparation, and so on). Computer algorithms are used to correlate the depth and resistance values to the dopant concentration at each level. The weak points

of SRP are the need for skilled operators, the sample preparation, and the destructive nature of the test. Its strong point is its ability to provide accurate dopant concentration measurements with no junction-depth limitation.

Unpatterned Surface Defects

Unpatterned wafers are bare silicon wafers or wafers with various blanket films used as test wafers and incorporated in process runs to provide a source of characterization information about process conditions. The unpatterned wafer may be polished to a mirror finish or have a film that has a rougher surface. After use in a process run, the unpatterned wafer is typically cleaned and reused, which adds cost to the manufacturing process and reduces a company's profit margin. Typical defects inspected for on unpatterned wafers used for process monitoring include particles, scratches, slip lines, and other material defects.

Defect-detection equipment for wafer-surface defects falls into two general categories: darkfield and brightfield optical detection.[15] *Brightfield detection* is the traditional light source for microscope equipment—it examines the wafer surface for defects with directly reflected visible light. With brightfield detection, horizontal surfaces reflect most of the light while slanted or vertical surfaces reflect less. *Darkfield detection* examines light scattered off defects located on the wafer surface. This is done by directing light to the wafer surface at a shallow angle through the outside of the optic's objective body (see Figure 7.15). This light impinges on the wafer surface and passes back up through the center of the optics. This action renders all flat surfaces black, while irregularities appear as bright lines. This fact makes darkfield detection useful for bringing out small defects on the wafer surface that might be difficult to see with brightfield detection. An example of darkfield detection is seeing dust particles in a ray of sunshine in a dark room. Both systems usually use some form of signal or image processing to locate defects based on the light signal received from the wafer surface.

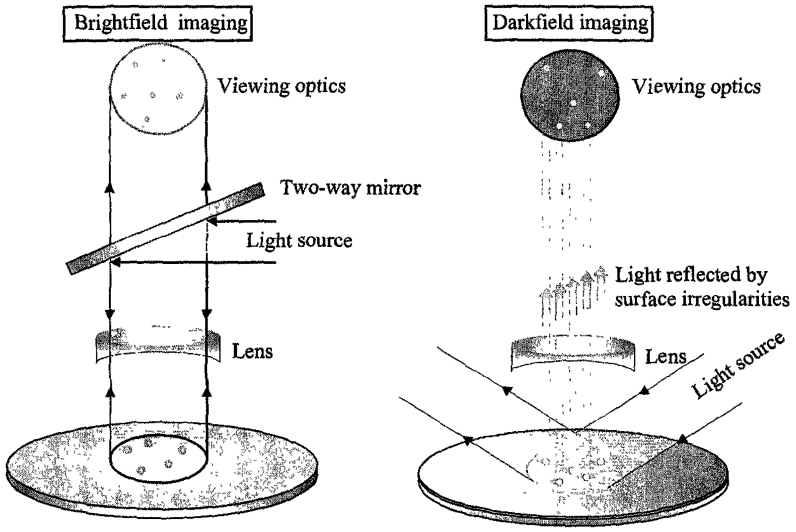


FIGURE 7.15 Darkfield and Brightfield Detection

Optical Microscopy ■ The *optical microscope,* or *light microscope,* has been used in science since the seventeenth century. For semiconductor manufacturing, it has traditionally been one of the most widely used methods of inspecting the wafer surface for defects such as particles and scratches. The optical microscope provides low-magnification views of the wafer, typically less than 1,000X magnification. Previously, optical microscopy was not capable of performing in the submicron regime. With shrinking device geometries, the size of killer defects to be detected has decreased, which has required the optical microscope to also improve over the years. Depending on the type of optical system used, particulate defects down to a size of 0.1 μm are currently able to be detected.[16] Not all layers on a chip are critical in terms of dimensions, and in many instances it is necessary to detect only gross defects. In this case, the optical microscope is quick and cost-effective.