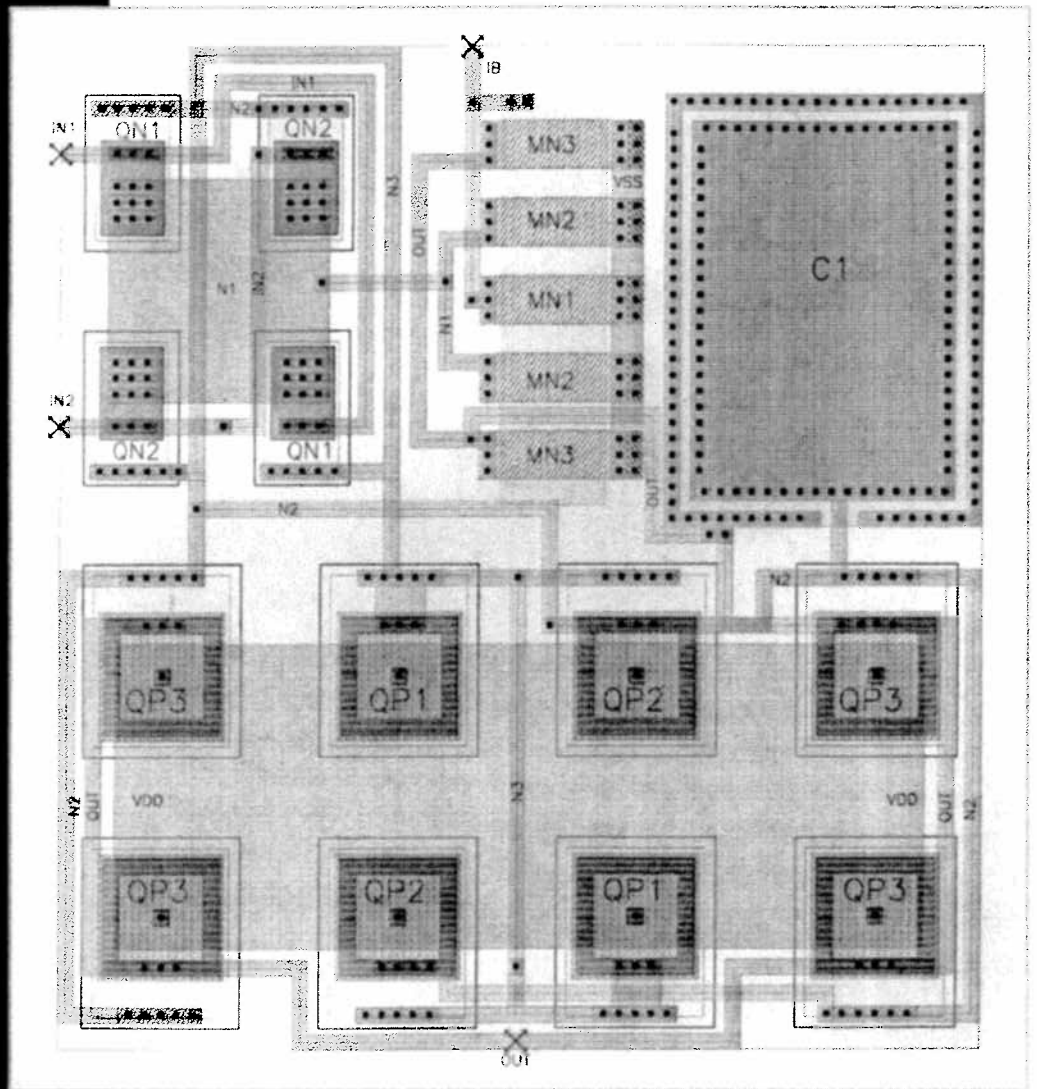


The Art of

ANALOG LAYOUT



Alan Hastings

The ART of
ANALOG LAYOUT

Alan Hastings



PRENTICE HALL
Upper Saddle River, NJ 07458

Library of Congress Cataloging-in-Publication Data

Hastings, Alan (Ray Alan)

The art of analog layout / Alan Hastings.

p. cm.

Includes bibliographical references and index.

ISBN 0-13-087061-7

1. Integrated circuits—Design and construction. 2. Layout (Printing) I. Title.

TK7874.H3926 2001

621.3815—dc21

00-045307

Vice president and editorial director, ECS: **Marcia Horton**

Publisher: **Tom Robbins**

Associate editor: **Alice Dworkin**

Editorial assistant: **Jessica Power**

Production editor: **Carlisle Communications, Ltd.**

Executive managing editor: **Vince O'Brien**

Managing editor: **David A. George**

Art director: **Jayne Conte**

Cover design: **Joseph Sengotta**

Art editor: **Adam Velthaus**

Manufacturing manager: **Trudy Piscioti**

Manufacturing buyer: **Dawn Murrin**

Assistant vice president of production and manufacturing, ESM: **David W. Riccardi**



Copyright © 2001 by Prentice-Hall, Inc.
Upper Saddle River, New Jersey 07458.

All rights reserved. No part of this book may be reproduced, in any form or by any means, without the permission in writing from the publisher.

The author and publisher of this book have used their best efforts in preparing this book. These efforts include the development, research, and testing of the theories to determine their effectiveness. The author and publisher make no warranty of any kind, expressed or implied, with regard to the documentation contained in this book.

Printed in the United States of America

10 9 8 7 6 5

ISBN 0-13-087061-7

Prentice-Hall International (UK) Limited, *London*

Prentice-Hall of Australia Pty. Limited, *Sydney*

Prentice-Hall Canada Inc., *Toronto*

Prentice-Hall Hispanoamericana, S.A., *Mexico*

Prentice-Hall of India Private Limited, *New Delhi*

Prentice-Hall of Japan, Inc., *Tokyo*

Pearson Education Asia Pte. Ltd., *Singapore*

Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

Contents

Preface xvii
Acknowledgments xix

1 *Device Physics*

- 1.1 Semiconductors 1
 - 1.1.1 Generation and Recombination 4
 - 1.1.2 Extrinsic Semiconductors 6
 - 1.1.3 Diffusion and Drift 9
- 1.2 PN Junctions 10
 - 1.2.1 Depletion Regions 10
 - 1.2.2 PN Diodes 13
 - 1.2.3 Schottky Diodes 15
 - 1.2.4 Zener Diodes 17
 - 1.2.5 Ohmic Contacts 19
- 1.3 Bipolar Junction Transistors 20
 - 1.3.1 Beta 22
 - 1.3.2 I-V Characteristics 23
- 1.4 MOS Transistors 24
 - 1.4.1 Threshold Voltage 27
 - 1.4.2 I-V Characteristics 29
- 1.5 JFET Transistors 31
- 1.6 Summary 33
- 1.7 Exercises 34

2 *Semiconductor Fabrication*

- 2.1 Silicon Manufacture 36
 - 2.1.1 Crystal Growth 37
 - 2.1.2 Wafer Manufacturing 38
 - 2.1.3 The Crystal Structure of Silicon 38
- 2.2 Photolithography 40
 - 2.2.1 Photoresists 40
 - 2.2.2 Photomasks and Reticles 41
 - 2.2.3 Patterning 42
- 2.3 Oxide Growth and Removal 42
 - 2.3.1 Oxide Growth and Deposition 43
 - 2.3.2 Oxide Removal 44
 - 2.3.3 Other Effects of Oxide Growth and Removal 46
 - 2.3.4 Local Oxidation of Silicon (LOCOS) 48

- 2.4 Diffusion and Ion Implantation 49
 - 2.4.1 Diffusion 50
 - 2.4.2 Other Effects of Diffusion 52
 - 2.4.3 Ion Implantation 53
- 2.5 Silicon Deposition 55
 - 2.5.1 Epitaxy 56
 - 2.5.2 Polysilicon Deposition 58
- 2.6 Metallization 58
 - 2.6.1 Deposition and Removal of Aluminum 59
 - 2.6.2 Refractory Barrier Metal 60
 - 2.6.3 Silicidation 62
 - 2.6.4 Interlevel Oxide, Interlevel Nitride, and Protective Overcoat 63
- 2.7 Assembly 64
 - 2.7.1 Mount and Bond 66
 - 2.7.2 Packaging 69
- 2.8 Summary 69
- 2.9 Exercises 69

3 *Representative Processes*

- 3.1 Standard Bipolar 72
 - 3.1.1 Essential Features 72
 - 3.1.2 Fabrication Sequence 73
 - Starting Material* 73
 - N-Buried Layer* 73
 - Epitaxial Growth* 74
 - Isolation Diffusion* 74
 - Deep-N+* 74
 - Base Implant* 75
 - Emitter Diffusion* 75
 - Contact* 76
 - Metallization* 76
 - Protective Overcoat* 77
 - 3.1.3 Available Devices 77
 - NPN Transistors* 77
 - PNP Transistors* 79
 - Resistors* 81
 - Capacitors* 83
 - 3.1.4 Process Extensions 84
 - Up-down Isolation* 84
 - Double-level Metal* 84
 - Schottky Diodes* 85
 - High-Sheet Resistors* 86
 - Super-beta Transistors* 86
- 3.2 Polysilicon-Gate CMOS 87
 - 3.2.1 Essential Features 88

3.2.2	Fabrication Sequence	89
	<i>Starting Material</i>	89
	<i>Epitaxial Growth</i>	89
	<i>N-well Diffusion</i>	89
	<i>Inverse Moat</i>	90
	<i>Channel Stop Implants</i>	90
	<i>LOCOS Processing and Dummy Gate Oxidation</i>	91
	<i>Threshold Adjust</i>	92
	<i>Polysilicon Deposition and Patterning</i>	93
	<i>Source/Drain Implants</i>	93
	<i>Contacts</i>	94
	<i>Metallization</i>	94
	<i>Protective Overcoat</i>	94
3.2.3	Available Devices	95
	<i>NMOS Transistors</i>	95
	<i>PMOS Transistors</i>	97
	<i>Substrate PNP Transistors</i>	98
	<i>Resistors</i>	98
	<i>Capacitors</i>	100
3.2.4	Process Extensions	100
	<i>Double-level Metal</i>	100
	<i>Silicidation</i>	101
	<i>Lightly Doped Drain (LDD) Transistors</i>	101
	<i>Extended-Drain, High-Voltage Transistors</i>	103
3.3	Analog BiCMOS	104
3.3.1	Essential Features	104
3.3.2	Fabrication Sequence	106
	<i>Starting Material</i>	106
	<i>N-buried Layer</i>	106
	<i>Epitaxial Growth</i>	106
	<i>N-well Diffusion and Deep-N+</i>	107
	<i>Base Implant</i>	107
	<i>Inverse Moat</i>	108
	<i>Channel Stop Implants</i>	108
	<i>LOCOS Processing and Dummy Gate Oxidation</i>	108
	<i>Threshold Adjust</i>	109
	<i>Polysilicon Deposition and Pattern</i>	109
	<i>Source/Drain Implants</i>	109
	<i>Metallization and Protective Overcoat</i>	110
	<i>Process Comparison</i>	110
3.3.3	Available Devices	111
	<i>NPN Transistors</i>	112
	<i>PNP Transistors</i>	112
	<i>Resistors</i>	115
3.4	Summary	115
3.5	Exercises	116

4 *Failure Mechanisms*

- 4.1 Electrical Overstress 118
 - 4.1.1 Electrostatic Discharge (ESD) 118
 - Effects* 120
 - Preventative Measures* 120
 - 4.1.2 Electromigration 121
 - Effects* 121
 - Preventative Measures* 122
 - 4.1.3 The Antenna Effect 122
- 4.2 Contamination 124
 - 4.2.1 Dry Corrosion 124
 - Effects* 124
 - Preventative Measures* 125
 - 4.2.2 Mobile Ion Contamination 125
 - Effects* 125
 - Preventative Measures* 126
- 4.3 Surface Effects 128
 - 4.3.1 Hot Carrier Injection 128
 - Effects* 128
 - Preventative Measures* 130
 - 4.3.2 Parasitic Channels and Charge Spreading 131
 - Effects* 131
 - Preventative Measures (Standard Bipolar)* 133
 - Preventative Measures (CMOS and BiCMOS)* 137
- 4.4 Parasitics 139
 - 4.4.1 Substrate Debiasing 140
 - Effects* 140
 - Preventative Measures* 142
 - 4.4.2 Minority-Carrier Injection 143
 - Effects* 143
 - Preventative Measures (Substrate Injection)* 146
 - Preventative Measures (Cross-injection)* 151
- 4.5 Summary 153
- 4.6 Exercises 153

5 *Resistors*

- 5.1 Resistivity and Sheet Resistance 156
- 5.2 Resistor Layout 158
- 5.3 Resistor Variability 162
 - 5.3.1 Process Variation 162
 - 5.3.2 Temperature Variation 163
 - 5.3.3 Nonlinearity 163
 - 5.3.4 Contact Resistance 166
- 5.4 Resistor Parasitics 167

- 5.5 Comparison of Available Resistors 170
 - 5.5.1 Base Resistors 170
 - 5.5.2 Emitter Resistors 171
 - 5.5.3 Base Pinch Resistors 172
 - 5.5.4 High-Sheet Resistors 173
 - 5.5.5 Epi Pinch Resistors 175
 - 5.5.6 Metal Resistors 176
 - 5.5.7 Poly Resistors 177
 - 5.5.8 NSD and PSD Resistors 180
 - 5.5.9 N-well Resistors 180
 - 5.5.10 Thin-film Resistors 181
- 5.6 Adjusting Resistor Values 182
 - 5.6.1 Tweaking Resistors 182
 - Sliding Contacts* 183
 - Sliding Heads* 184
 - Trombone Slides* 184
 - Metal Options* 184
 - 5.6.2 Trimming Resistors 185
 - Fuses* 185
 - Zener Zaps* 189
 - Laser Trims* 190
- 5.7 Summary 191
- 5.8 Exercises 192

6 *Capacitors*

- 6.1 Capacitance 194
- 6.2 Capacitor Variability 200
 - 6.2.1 Process Variation 200
 - 6.2.2 Voltage Modulation and Temperature Variation 201
- 6.3 Capacitor Parasitics 203
- 6.4 Comparison of Available Capacitors 205
 - 6.4.1 Base-emitter Junction Capacitors 205
 - 6.4.2 MOS Capacitors 207
 - 6.4.3 Poly-poly Capacitors 209
 - 6.4.4 Miscellaneous Styles of Capacitors 211
- 6.5 Summary 212
- 6.6 Exercises 212

7 *Matching of Resistors and Capacitors*

- 7.1 Measuring Mismatch 214
- 7.2 Causes of Mismatch 217
 - 7.2.1 Random Statistical Fluctuations 217
 - 7.2.2 Process Biases 219
 - 7.2.3 Pattern Shift 220

- 7.2.4 Variations in Polysilicon Etch Rate 222
- 7.2.5 Diffusion Interactions 224
- 7.2.6 Stress Gradients and Package Shifts 226
 - Piezoresistivity* 227
 - Gradients and Centroids* 229
 - Common-centroid Layout* 231
 - Location and Orientation* 235
- 7.2.7 Temperature Gradients and Thermoelectrics 236
 - Thermal Gradients* 238
 - Thermoelectric Effects* 240
- 7.2.8 Electrostatic Interactions 242
 - Voltage Modulation* 242
 - Charge Spreading* 245
 - Dielectric Polarization* 246
 - Dielectric Relaxation* 248
- 7.3 Rules for Device Matching 249
 - 7.3.1 Rules for Resistor Matching 249
 - 7.3.2 Rules for Capacitor Matching 253
- 7.4 Summary 257
- 7.5 Exercises 257

8 *Bipolar Transistors*

- 8.1 Topics in Bipolar Transistor Operation 260
 - 8.1.1 Beta Rolloff 262
 - 8.1.2 Avalanche Breakdown 262
 - 8.1.3 Thermal Runaway and Secondary Breakdown 264
 - 8.1.4 Saturation in NPN Transistors 266
 - 8.1.5 Saturation in Lateral PNP Transistors 270
 - 8.1.6 Parasitics of Bipolar Transistors 272
- 8.2 Standard Bipolar Small-signal Transistors 274
 - 8.2.1 The Standard Bipolar NPN Transistor 274
 - Construction of Small-signal NPN Transistors* 276
 - 8.2.2 The Standard Bipolar Substrate PNP Transistor 279
 - Construction of Small-signal Substrate PNP Transistors* 281
 - 8.2.3 The Standard Bipolar Lateral PNP Transistor 283
 - Construction of Small-signal Lateral PNP Transistors* 285
 - 8.2.4 High-voltage Bipolar Transistors 291
- 8.3. Alternative Small-signal Bipolar Transistors 293
 - 8.3.1 Extensions to Standard Bipolar 293
 - 8.3.2 Analog BiCMOS Bipolar Transistors 294
 - 8.3.3 Bipolar Transistors in a CMOS Process 297
 - 8.3.4 Advanced-technology Bipolar Transistors 299
- 8.4 Summary 302
- 8.5 Exercises 303

9 Applications of Bipolar Transistors

- 9.1 Power Bipolar Transistors 306
 - 9.1.1 Failure Mechanisms of NPN Power Transistors 307
 - Emitter Debiasing* 307
 - Thermal Runaway and Secondary Breakdown* 309
 - 9.1.2 Layout of Power NPN Transistors 311
 - The Interdigitated-emitter Transistor* 311
 - The Wide-emitter Narrow-contact Transistor* 314
 - The Christmas-tree Device* 315
 - The Cruciform-emitter Transistor* 316
 - Power Transistor Layout in Analog BiCMOS* 317
 - Selecting a Power Transistor Layout* 318
 - 9.1.3 Saturation Detection and Limiting 319
- 9.2 Matching Bipolar Transistors 322
 - 9.2.1 Random Variations 323
 - 9.2.2 Emitter Degeneration 325
 - 9.2.3 NBL Shadow 327
 - 9.2.4 Thermal Gradients 328
 - 9.2.5 Stress Gradients 332
- 9.3 Rules for Bipolar Transistor Matching 334
 - 9.3.1 Rules for Matching NPN Transistors 335
 - 9.3.2 Rules for Matching Lateral PNP Transistors 337
- 9.4 Summary 340
- 9.5 Exercises 340

10 Diodes

- 10.1 Diodes in Standard Bipolar 343
 - 10.1.1 Diode-connected Transistors 343
 - 10.1.2 Zener Diodes 346
 - Surface Zener Diodes* 347
 - Buried Zeners* 349
 - 10.1.3 Schottky Diodes 352
- 10.2 Diodes in CMOS and BiCMOS Processes 356
- 10.3 Matching Diodes 359
 - 10.3.1 Matching PN Junction Diodes 359
 - 10.3.2 Matching Zener Diodes 360
 - 10.3.3 Matching Schottky Diodes 361
- 10.4 Summary 362
- 10.5 Exercises 362

11 MOS Transistors

- 11.1 Topics in MOS Transistor Operation 364
 - 11.1.1 Modeling the MOS Transistor 364
 - Device Transconductance* 365
 - Threshold Voltage* 367

11.1.2	Parasitics of MOS Transistors	370
	<i>Breakdown Mechanisms</i>	372
	<i>CMOS Latchup</i>	375
11.2	Self-aligned Poly-Gate CMOS Transistors	376
11.2.1	Coding the MOS Transistor	377
	Width and Length	378
11.2.2	N-well and P-well Processes	379
11.2.3	Channel Stops	381
11.2.4	Threshold Adjust Implants	383
11.2.5	Scaling the Transistor	386
11.2.6	Variant Structures	388
	<i>Serpentine Transistors</i>	391
	<i>Annular Transistors</i>	391
11.2.7	Backgate Contacts	393
11.3	Summary	396
11.4	Exercises	396

12 *Applications of MOS Transistors*

12.1	Extended-voltage Transistors	399
12.1.1	LDD and DDD Transistors	400
12.1.2	Extended-drain Transistors	403
	<i>Extended-drain NMOS Transistors</i>	403
	<i>Extended-drain PMOS Transistors</i>	405
12.1.3	Multiple Gate Oxides	405
12.2	Power MOS Transistors	407
	<i>Thermal Runaway</i>	407
	<i>Secondary Breakdown</i>	408
	<i>Rapid Transient Overload</i>	408
	<i>MOS Switches versus Bipolar Switches</i>	409
12.2.1	Conventional MOS Power Transistors	410
	<i>The Rectangular Device</i>	411
	<i>The Diagonal Device</i>	413
	<i>Computation of R_M</i>	413
	<i>Other Considerations</i>	414
	<i>Nonconventional Structures</i>	416
12.2.2	DMOS Transistors	417
	<i>The Lateral DMOS Transistor</i>	418
	<i>The DMOS NPN</i>	420
12.3	The JFET Transistor	422
12.3.1	Modeling the JFET	422
12.3.2	JFET Layout	423
12.4	MOS Transistor Matching	426
12.4.1	Geometric Effects	427
	<i>Gate Area</i>	428
	<i>Gate Oxide Thickness</i>	428

- Channel Length Modulation* 429
- Orientation* 429
- 12.4.2 Diffusion and Etch Effects 430
 - Polysilicon Etch Rate Variations* 430
 - Contacts Over Active Gate* 431
 - Diffusions Near the Channel* 432
 - PMOS versus NMOS Transistors* 432
- 12.4.3 Thermal and Stress Effects 433
 - Oxide Thickness Gradients* 433
 - Stress Gradients* 433
 - Metallization-induced Stresses* 434
 - Thermal Gradients* 434
- 12.4.4 Common-centroid Layout of MOS Transistors 435
- 12.5 Rules for MOS Transistor Matching 439
- 12.6 Summary 442
- 12.7 Exercises 443

13 *Special Topics*

- 13.1 Merged Devices 445
 - 13.1.1 Flawed Device Mergers 446
 - 13.1.2 Successful Device Mergers 450
 - 13.1.3 Low-risk Merged Devices 452
 - 13.1.4 Medium-risk Merged Devices 453
 - 13.1.5 Devising New Merged Devices 455
- 13.2 Guard Rings 455
 - 13.2.1 Standard Bipolar Electron Guard Rings 456
 - 13.2.2 Standard Bipolar Hole Guard Rings 457
 - 13.2.3 Guard Rings in CMOS and BiCMOS Designs 458
- 13.3 Single-level Interconnection 460
 - 13.3.1 Mock Layouts and Stick Diagrams 461
 - 13.3.2 Techniques for Crossing Leads 463
 - 13.3.3 Types of Tunnels 464
- 13.4 Constructing the Pading 466
 - 13.4.1 Scribe Streets and Alignment Markers 466
 - 13.4.2 Bondpads, Trimpads, and Testpads 468
 - 13.4.3 ESD Structures 471
 - Zener Clamp* 473
 - Two-stage Zener Clamps* 475
 - Buffered Zener Clamp* 476
 - V_{CES} Clamp* 478
 - V_{CES} Clamp* 479
 - Antiparallel Diode Clamps* 480
 - Additional ESD Structures for CMOS Processes* 480
 - 13.4.4 Selecting ESD Structures 483
- 13.5 Exercises 485

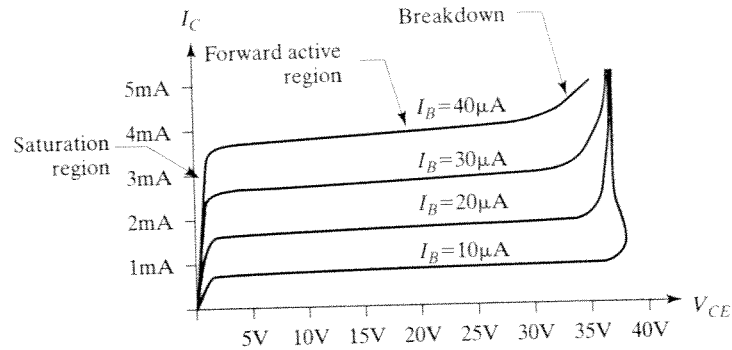
14 *Assembling the Die*

- 14.1 Die Planning 488
 - 14.1.1 Cell Area Estimation 489
 - Resistors* 489
 - Capacitors* 489
 - Vertical Bipolar Transistors* 489
 - Lateral PNP Transistors* 490
 - MOS Transistors* 490
 - MOS Power Transistors* 490
 - Computing Cell Area* 491
 - 14.1.2 Die Area Estimation 491
 - 14.1.3 Gross Profit Margin 494
- 14.2 Floorplanning 495
- 14.3 Top-level Interconnection 500
 - 14.3.1 Principles of Channel Routing 501
 - 14.3.2 Special Routing Techniques 503
 - Kelvin Connections* 503
 - Noisy Signals and Sensitive Signals* 504
 - 14.3.3 Electromigration 506
 - 14.3.4 Minimizing Stress Effects 508
- 14.4 Conclusion 510
- 14.5 Exercises 510

Appendices

- A. Table of Acronyms Used in the Text 513
- B. The Miller Indices of a Cubic Crystal 516
- C. Sample Layout Rules 519
- D. Mathematical Derivations 527
- E. Sources for Layout Editor Software 532

Index 533

FIGURE 1.21 Typical I-V plot of an NPN transistor.

designers because the forward biasing of the collector-base junction injects minority carriers into the neutral collector. Section 8.1.4 discusses the effects of saturation upon integrated bipolar transistors in greater detail.

The collector-emitter voltage in the forward active region is large enough to reverse bias the collector-base junction. Ohmic drops in the collector no longer significantly reduce the electric field across the collector-base junction, so the current flow through the transistor now depends solely upon beta. The slight upward tilt to the current curves results from the *Early effect*. As the reverse bias on the collector-base junction increases, the depletion region at this junction widens and consequently the neutral base narrows. Since beta depends on base width, it increases slightly as the collector-emitter voltage rises. The Early effect can be minimized by using a very lightly doped collector, so the depletion region extends primarily into the collector rather than into the base.

Beyond a certain collector-emitter voltage, the collector current increases rapidly. This effect limits the maximum operating voltage of the transistor. In the case of a typical integrated NPN transistor, this voltage equals some 30V–40V. The increased current flow results from either one of two effects, the first of which is avalanche breakdown. The collector-base junction will avalanche if it is sufficiently reverse-biased. A wide lightly doped collector region can greatly increase the avalanche voltage rating, and discrete power transistors can achieve operating voltages of more than a thousand volts.

The second limiting mechanism is *base punchthrough*. Punchthrough occurs when the collector-base depletion region reaches all the way through the base and merges with the base-emitter depletion region. Once this occurs, carriers can flow directly from emitter to collector, and current is limited only by the resistance of the neutral collector and emitter. The resulting rapid increase in collector current mimics the effects of avalanche breakdown.

Base punchthrough is often observed in high-gain transistors. For example, *super-beta* transistors use an extremely thin base region to obtain betas of a thousand or more. Base punchthrough limits the operating voltage of these devices to a couple of volts. Super-beta transistors also display a pronounced Early effect because of the encroachment of the collector-base depletion region into the extremely thin neutral base. General-purpose transistors use wider base regions to reduce the Early effect, and their operating voltages are usually limited by avalanche instead of base punchthrough (Section 8.1.2).

1.4 MOS TRANSISTORS

The bipolar junction transistor amplifies a small change in input current to provide a large change in output current. The gain of a bipolar transistor is thus defined as the ratio of output to input current (beta). Another type of transistor, called a *field-*

effect transistor (FET), transforms a change in input voltage into a change in output current. The gain of an FET is measured by its *transconductance*, defined as the ratio of change in output current to change in input voltage.

The field-effect transistor is so named because its input terminal (called its *gate*) influences the flow of current through the transistor by projecting an electric field across an insulating layer. Virtually no current flows through this insulator, so the gate current of a FET transistor is vanishingly small. The most common type of FET uses a thin silicon dioxide layer as an insulator beneath the gate electrode. This type of transistor is called a *metal-oxide-semiconductor* (MOS) transistor, or alternatively, a *metal-oxide-semiconductor field-effect transistor* (MOSFET). MOS transistors have replaced bipolars in many applications because they are smaller and can often operate using less power.

The MOS transistor can be better understood by first considering a simpler device called a *MOS capacitor*. This device consists of two electrodes, one of metal and one of extrinsic silicon, separated by a thin layer of silicon dioxide (Figure 1.22A). The metal electrode forms the *gate*, while the semiconductor slab forms the *backgate* or *body*. The insulating oxide layer between the two is called the *gate dielectric*. The illustrated device has a backgate consisting of lightly doped P-type silicon. The electrical behavior of this MOS capacitor can be demonstrated by grounding the backgate and biasing the gate to various voltages. The MOS capacitor of Figure 1.22A has a gate potential of 0V. The difference in work functions between the metal gate and the semiconductor backgate causes a small electric field to appear across the dielectric. In the illustrated device, this field biases the metal plate slightly positive with respect to the P-type silicon. This electric field attracts electrons from deep within the silicon up toward the surface, while it repels holes away from the surface. The field is weak, so the change in carrier concentrations is small and the overall effect upon the device characteristics is minimal.

Figure 1.22B shows what occurs when the gate of the MOS capacitor is biased positively with respect to the backgate. The electric field across the gate dielectric strengthens and more electrons are drawn up from the bulk. Simultaneously, holes are repelled away from the surface. As the gate voltage rises, a point is reached where more electrons than holes are present at the surface. Due to the excess electrons, the surface layers of the silicon behave as if they were N-type. The apparent reversal of doping polarity is called *inversion* and the layer of silicon that inverts is called a *channel*. As the gate voltage increases still further, more electrons accumulate at the surface and the channel becomes more strongly inverted. The voltage at which the channel just begins to form is called the *threshold voltage* V_t . When the voltage difference between gate and backgate is less than the threshold voltage, no channel forms. When the voltage difference exceeds the threshold voltage, a channel appears.

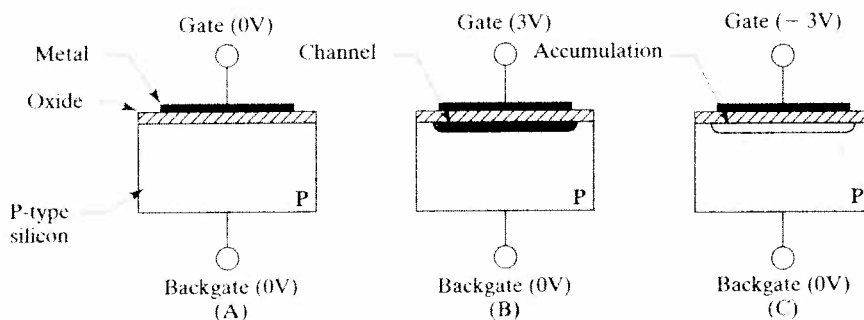
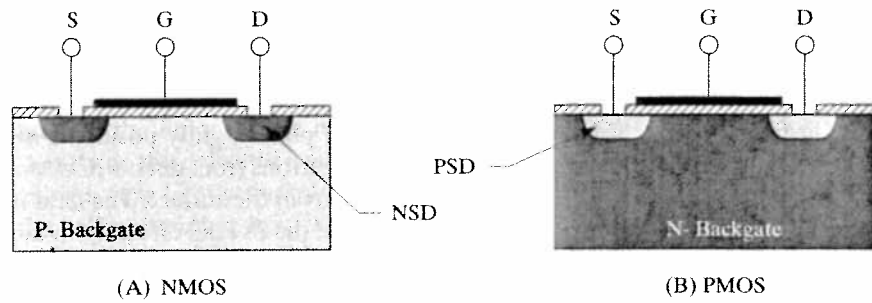


FIGURE 1.22 MOS capacitor: (A) unbiased ($V_{BG} = 0V$), (B) inversion ($V_{BG} = 3V$), (C) accumulation ($V_{BG} = -3V$).

Figure 1.22C shows what happens if the gate of the MOS capacitor is biased negatively with respect to the backgate. The electric field now reverses, drawing holes toward the surface and repelling electrons away from it. The surface layers of silicon appear to be more heavily doped, and the device is said to be in *accumulation*.

The behavior of the MOS capacitor can be utilized to form a true MOS transistor. Figure 1.23A shows the cross section of the resulting device. The gate, dielectric, and backgate remain as before. Two additional regions are formed by selectively doping the silicon on either side of the gate. One of these regions is called the *source* and the other is called the *drain*. Imagine that the source and backgate are both grounded and that a positive voltage is applied to the drain. As long as the gate-to-backgate voltage remains less than the threshold voltage, no channel forms. The PN junction formed between drain and backgate is reverse-biased, so very little current flows from drain to backgate. If the gate voltage exceeds the threshold voltage, a channel forms beneath the gate dielectric. This channel acts like a thin film of N-type silicon shorting the source to the drain. A current consisting of electrons flows from the source across the channel to the drain. In summary, drain current will only flow if the gate-to-source voltage V_{GS} exceeds the threshold voltage V_t .

FIGURE 1.23 Cross sections of MOSFET transistors: NMOS (A) and PMOS (B). In these diagrams, S = Source, G = Gate, and D = Drain. The backgate connections, though present, are not illustrated.



The source and drain of a MOS transistor are interchangeable, as both are simply N-type regions formed in the P-type backgate. In many cases, these two regions are identical and the terminals can be reversed without changing the behavior of the device. Such a device is said to be *symmetric*. In a symmetric MOS transistor the labeling of source and drain becomes somewhat arbitrary. By definition, carriers flow out of the source and into the drain. The identity of the source and the drain therefore depends on the biasing of the device. Sometimes the bias applied across the transistor fluctuates and the two terminals swap roles. In such cases, the circuit designer must arbitrarily designate one terminal the drain and the other the source.

Asymmetric MOS transistors are designed with different source and drain dopings and geometries. There are several reasons why transistors may be made asymmetric, but the result is the same in every case. One terminal is optimized to function as the drain and the other as the source. If source and drain are swapped, then the performance of the device will suffer.

The transistor depicted in Figure 1.23A has an N-type channel and is therefore called an *N-channel MOS transistor*, or NMOS. *P-channel MOS (PMOS)* transistors also exist. Figure 1.23B shows a sample PMOS transistor consisting of a lightly doped N-type backgate with P-type source and drain regions. If the gate of this transistor is biased positive with respect to the backgate, then electrons are drawn to the surface and holes are repelled away from it. The surface of the silicon accumulates, and no channel forms. If the gate is biased negative with respect to the backgate, then holes are drawn to the surface, and a channel forms. The PMOS transistor thus

has a negative threshold voltage. Engineers often ignore the sign of the threshold voltage since it is normally positive for NMOS transistors and negative for PMOS transistors. An engineer might say, "The PMOS V_t has increased from 0.6V to 0.7V" when in actuality the PMOS V_t has shifted from -0.6V to -0.7V.

1.4.1. Threshold Voltage

The *threshold voltage* of a MOS transistor equals the gate-to-source bias required to just form a channel with the backgate of the transistor connected to the source. If the gate-to-source bias is less than the threshold voltage, then no channel forms. The threshold voltage exhibited by a given transistor depends on a number of factors, including backgate doping, dielectric thickness, gate material, and excess charge in the dielectric. Each of these effects will be briefly examined.

Backgate doping has a major effect on the threshold voltage. If the backgate is doped more heavily, then it becomes more difficult to invert. A stronger electric field is required to achieve inversion, and the threshold voltage increases. The backgate doping of an MOS transistor can be adjusted by performing a shallow implant beneath the surface of the gate dielectric to dope the channel region. This type of implant is called a *threshold adjust implant* (or V_t *adjust implant*).

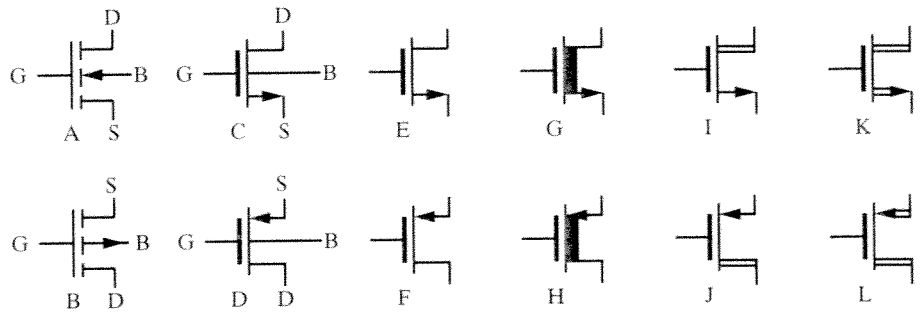
Consider the effects of a V_t adjust implant upon an NMOS transistor. If the implant consists of acceptors, then the silicon surface becomes more difficult to invert and the threshold voltage increases. If the implant consists of donors, then the surface becomes easier to invert and the threshold decreases. If enough donors are implanted, the surface of the silicon can actually become counterdoped. In this case, a thin layer of N-type silicon forms a permanent channel at zero gate bias. The channel becomes more strongly inverted as the gate bias increases. As the gate bias is decreased, the channel becomes less strongly inverted and at some point it vanishes. The threshold voltage of this NMOS transistor is actually negative. Such a transistor is called a *depletion-mode NMOS*, or simply a *depletion NMOS*. In contrast, an NMOS with a positive threshold voltage is called an *enhancement-mode NMOS*, or *enhancement NMOS*. The vast majority of commercially fabricated MOS transistors are enhancement-mode devices, but there are a few applications that require depletion-mode devices. A depletion-mode PMOS can also be constructed. Such a device will have a positive threshold voltage.

Depletion-mode devices should always be explicitly identified as such. One cannot rely on the sign of the threshold voltage to convey this information, because many engineers customarily ignore threshold polarities. Therefore, one should say "a depletion-mode PMOS with a threshold of 0.7V," rather than a PMOS with a threshold of 0.7V." Many engineers would interpret the latter statement as indicating an enhancement PMOS with a threshold of -0.7V rather than a depletion PMOS with a threshold of +0.7V. Explicitly referring to depletion-mode devices as such eliminates any possibility of confusion.

Special symbols are often used to distinguish between different types of MOS transistors. Figure 1.24 shows a representative collection of these symbols.¹⁰ Symbols A and B are the standard symbols for NMOS and PMOS transistors, respectively. These symbols are not commonly used in the industry; instead symbols

¹⁰ Symbols A, B, E, F, G, and H are used by various authors; see A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design* (New York: John Wiley and Sons, 1984), pp. 112-113; also P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed. (New York: John Wiley and Sons, 1993), p. 60. The *J. Solid State Circuits* also uses three-terminal MOS symbols but differentiates PMOS devices by placing a bubble on their gate leads.

FIGURE 1.24 MOSFET symbols: A, B: standard symbols; C, D: industry symbols (four-terminal); E, F: industry symbols (three-terminal); G, H: depletion-mode devices; I, J: asymmetric high-voltage MOS symbols; K, L: symmetric high-voltage MOS symbols.



C and D are preferred for NMOS and PMOS transistors, respectively. These symbols intentionally resemble NPN and PNP transistors. This convention helps highlight the essential similarities between MOS and bipolar circuits. Symbols E and F are sometimes employed when the backgates of the transistors connect to known potentials. Every MOS transistor has a backgate, so this terminal must always connect to something. Symbols E and F are potentially confusing, because the reader must infer the backgate connections. These symbols are nonetheless very popular because they make schematics much more legible. Symbols G and H are often used for depletion-mode devices, where the solid bar from drain to source represents the channel present at zero bias. Symbols I and J are sometimes employed for asymmetric transistors with high-voltage drains, and symbols K and L are used for symmetric transistors with high-voltage terminations for both source and drain. There are many other schematic symbols for MOS transistors; the ones shown in Figure 1.24 form only a representative sample.

Returning to the discussion of threshold voltage, the dielectric also plays an important role in determining the threshold voltage. A thicker dielectric weakens the electric field by separating the charges by a greater distance. Thus, thicker dielectrics increase the threshold voltage while thinner ones reduce it. In theory, the material of the dielectric also affects the electric field strength. In practice, almost all MOS transistors use pure silicon dioxide as the gate dielectric. This material can be grown in extremely thin films of exceptional purity and uniformity; no other material has comparable properties. Alternate dielectric materials therefore have very limited application.¹¹

The gate electrode material also affects the threshold voltage of the transistor. As mentioned above, an electric field appears across the gate oxide when the gate and backgate are shorted together. This field is produced by the difference in work functions between the gate and backgate materials. Most practical transistors use heavily doped polysilicon for the gate electrode. The work function of polysilicon can be varied to a limited degree by changing its doping.

A potentially troublesome source of threshold voltage variation comes from the presence of excess charges in the gate oxide or along the interfaces between the oxide and the silicon surface. These charges may consist of ionized impurity atoms, trapped carriers, or structural defects. The presence of trapped electric charge in the dielectric or along its interfaces alters the electric field and therefore the threshold voltage. If the amount of trapped charge varies with time, temperature, or applied bias, then the threshold voltage will also vary. This subject is discussed in greater detail in Section 4.2.2.

¹¹ A few devices have been fabricated using high-permittivity materials such as silicon nitride for the gate dielectric. Some authors use the term *insulated-gate field effect transistor* (IGFET) to refer to all MOS-like transistors, including those with non-oxide dielectrics.

1.4.2. I-V Characteristics

The performance of an MOS transistor can be graphically illustrated by drawing a family of I-V curves similar to those used for bipolar transistors. Figure 1.25 shows a typical set of curves for an enhancement NMOS. The source and backgate were connected together to obtain these particular curves. The vertical axis measures drain current I_D , while the horizontal axis measures drain-to-source voltage V_{DS} . Each curve represents a specific gate-to-source voltage V_{GS} . The general character of the curves resembles that of the bipolar transistor shown in Figure 1.21, but the family of curves for an MOS transistor are obtained by stepping gate voltage, while those for a bipolar transistor are obtained by stepping base current.

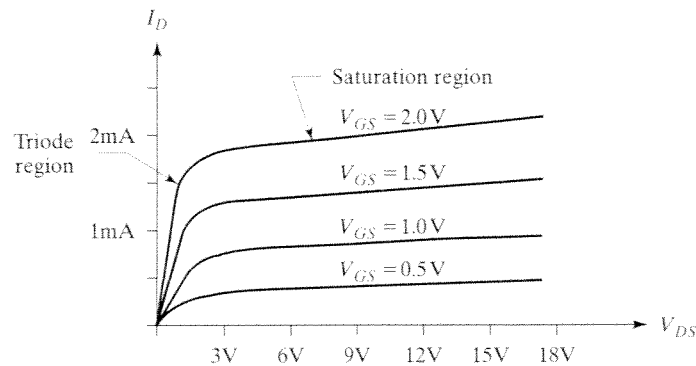


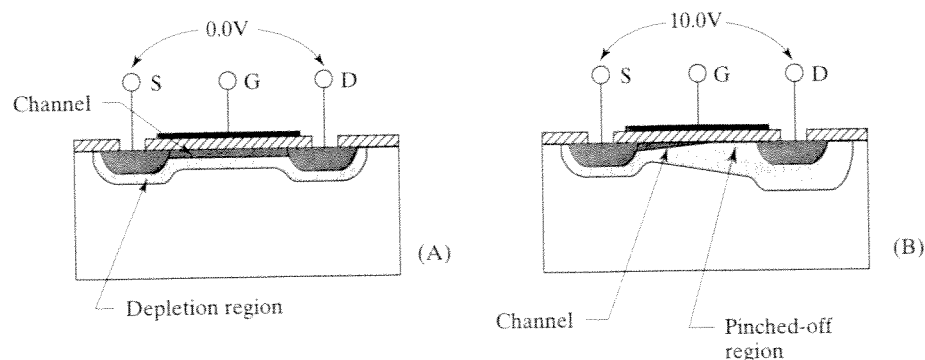
FIGURE 1.25 Typical I-V plot of an NMOS transistor.

At low drain-to-source voltages the MOS channel behaves resistively, and the drain current increases linearly with voltage. This region of operation is called the *linear region* or *triode region*. This roughly corresponds to the saturation region of a bipolar transistor. The drain current levels off to an approximately constant value when the drain-to-source voltage exceeds the difference between the gate-to-source voltage and the threshold voltage. This region is called the *saturation region*, and it roughly corresponds to the forward active region of a bipolar transistor. The term *saturation* thus has very different meanings for MOS and bipolar transistors.

The behavior of the MOS transistor in the linear region is easily explained. The channel acts as a film of doped silicon with a characteristic resistance that depends upon the carrier concentration. The current increases linearly with voltage, exactly as one would expect of a resistor. Higher gate voltages produce larger carrier concentrations and therefore lessen the resistance of the channel. PMOS transistors behave similarly to NMOS transistors, but since holes have lower mobilities than electrons, the apparent resistance of the channel is considerably greater. The effective resistance of an MOS transistor operating in the triode region is symbolized $R_{DS(on)}$.

MOS transistors saturate because of a phenomenon called *pinch-off*. While the drain-to-source voltage remains small, a depletion region of uniform thickness surrounds the channel (Figure 1.26A). As the drain becomes more positive with respect to the source, the depletion region begins to thicken at the drain end. This depletion region intrudes into the channel and narrows it. Eventually the channel depletes all the way through and it is said to have *pinched off* (Figure 1.26B). Carriers move down the channel propelled by the relatively weak electric field along it. When they reach the edge of the pinched-off region, they are sucked across the depletion region by the strong electric field. The voltage drop across the channel does not increase as the drain voltage is increased; instead the pinched-off region widens. Thus, the drain current reaches a limit and ceases to increase.

FIGURE 1.26 Behavior of a MOS transistor under bias: (A) $V_{DS} = 0V$ (triode region); (B) $V_{DS} = 10V$ (saturation region).



The drain current curves actually tilt slightly upward in the saturation region. This tilt is caused by *channel length modulation*, which is the MOS equivalent of the Early effect. Increases in drain voltage cause the pinched-off region to widen and the channel length to shorten. The shorter channel still has the same potential drop across it, so the electric field intensifies and the carriers move more rapidly. The drain current thus increases slightly with increasing drain-to-source voltage.

The I-V curves of Figure 1.25 were obtained with the backgate of the transistor connected to the source. If the backgate is biased independently of the source, then the apparent threshold voltage of the transistor will vary. If the source of an NMOS transistor is biased above its backgate, then its apparent threshold voltage increases. If the source of a PMOS transistor is biased below its backgate, then its threshold voltage decreases (it becomes more negative). This *backgate effect*, or *body effect*, arises because the backgate-to-source voltage modulates the depletion region beneath the channel. This depletion region widens as the backgate-to-source differential increases, and it intrudes into the channel as well as into the backgate. A high backgate-to-source differential will thin the channel, which in turn raises the apparent threshold voltage. The intrusion of the depletion region into the channel becomes more significant as the backgate doping rises, and this in turn increases the magnitude of the body effect.

MOS transistors are normally considered majority carrier devices, which conduct only after a channel forms. This simplistic view does not explain the low levels of conduction that occur at gate-to-source voltages just less than the threshold voltage. The formation of a channel is a gradual process. As the gate-to-source voltage increases, the gate first attracts small numbers of minority carriers to the surface. The concentration of minority carriers rises as the voltage increases. When the gate-to-source voltage exceeds the threshold, the number of minority carriers becomes so large that the surface of the silicon inverts and a channel forms. Before this occurs, minority carriers can still move from the source to the drain by diffusion. This *subthreshold conduction* produces currents that are much smaller than those that would flow if a channel were present. However, they are still many orders of magnitude greater than junction leakages. Subthreshold conduction is typically significant only when the gate-to-source voltage is within about 0.3V of the threshold voltage. This is sufficient to cause serious "leakage" problems in low- V_t devices. Some electrical circuits actually take advantage of the exponential voltage-to-current relationship of subthreshold conduction, but these circuits cannot operate at temperatures much in excess of 100°C because the junction leakages become so large that they overwhelm the tiny subthreshold currents.

As with bipolar transistors, MOS transistors can break down by either avalanche or punchthrough. If the voltage across the depletion region at the drain becomes so large that avalanche multiplication occurs, the drain current increases rapidly. Similarly, if the entire channel pinches off, then the source and drain will be shorted by the resulting depletion region and the transistor will punch through.

The operating voltage of an MOS transistor is often limited to a value considerably below the onset of avalanche or punchthrough by a long-term degradation mechanism called *hot carrier injection*. Carriers that traverse the pinched-off portion of the drain are accelerated by the strong electric field present here. The carriers can achieve velocities far beyond those normally associated with room-temperature thermal diffusion, so they are called *hot carriers*. When these carriers collide with atoms near the silicon surface, some of them are deflected up into the gate oxide, and a few of these become trapped. Slowly, over a long period of operation, the concentration of these trapped carriers increases and the threshold voltage shifts. Hot hole injection occurs less readily than hot electron injection because the lower mobility of holes limits their velocity and therefore their ability to surmount the oxide interface. For this reason, NMOS transistors are frequently limited to lower operating voltages than PMOS transistors of similar construction. Various techniques have been devised to limit hot carrier injection (Section 12.1).

1.5 JFET TRANSISTORS

The MOS transistor represents only one type of field-effect transistor. Another is the *junction field-effect transistor* or JFET. This device uses the depletion regions surrounding reverse-biased junctions as a gate dielectric. Figure 1.27A shows a cross-section of an N-channel JFET. This device consists of a bar of lightly doped N-type silicon called the *body* into which two P-type diffusions have been driven from opposite sides. The thin region of N-type silicon remaining between the junctions forms the *channel* of the JFET. The two diffusions act as the *gate* and the *backgate* and the opposite ends of the body form the *source* and the *drain*.

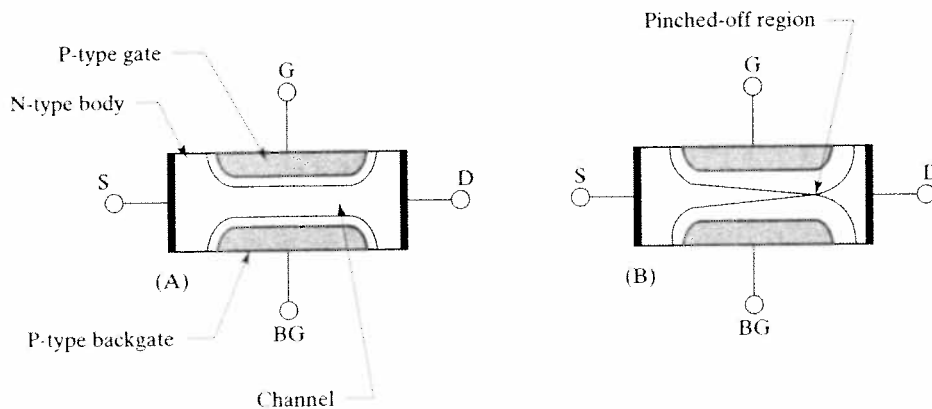


FIGURE 1.27 Cross sections of an N-channel JFET transistor operating in the linear region (A) and in saturation (B). In both diagrams, S = Source, D = Drain, G = Gate, and BG = Backgate.

Suppose that all four terminals of the N-channel JFET are grounded. Depletion regions form around the gate-body and backgate-body junctions. These depletion regions extend into the lightly doped channel, but they do not actually touch one another. A channel therefore exists from the drain to the source. If the drain voltage rises above the source voltage, then a current flows through the channel from